# Determining the Geographical Origin of Cocoa Beans in Chocolate using Stable Isotope Ratios

Tara Lee

A thesis submitted for the degree of
Master of Science
University of Otago, Dunedin, New Zealand

November 2020

# Abstract

Sustainability of cocoa bean farms is becoming an increasingly important matter. With more chocolate showing more sustainability labels, there needs to be confidence in the consumer's mind on these labels. One way in which to build this confidence is to ensure traceability of the cocoa bean from farm to the final chocolate product.

The stable isotope ratio of cocoa bean samples were measured using IRMS to see if their isotope ratio profile reflects their country of origin. $\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$ were all measured on the bulk unshelled cocoa bean, along with $\delta^{2}H$ and $\delta^{13}C$ of the three most abundant fatty acids within cocoa beans; Palmitic acid, Stearic acid and Oleic acid. The results were analysed using PCA and showed that the samples tended to split into two groups. Nicaragua, Ghana and Vietnam and Papua New Guinea, The Solomon Islands and Samoa.

Two chocolate samples were made from cocoa beans which originated from Ghana. The $\delta^{13}C$ and $\delta^{2}H$ of extracted fatty acids were analysed as their methyl esters (FAMEs) and compared to the results of the cocoa bean fatty acids. These results showed that the chocolate samples had isotope ratio values closer to that of Vietnam than Ghana. However, the samples from Ghana were also found to be close to those of Vietnam. As an investigative study, this study showed that the method of FAME isotope ratios could potentially be used in future to determine origin of cocoa beans. With some more samples and research, this could lead to the possibility of using it on cocoa beans within a mixture to determine origin.

This study also investigated the gas chromatography stable isotope analysis of underivatized fatty acids extracted from cocoa beans samples. This method is a way to analyse the fatty acids without the need to conduct mass balance calculations needed for FAME analysis. Although the chromatographic responses were found to be much lower than that of the FAME, results seen from one chromatographic peak showed similar trends observed for the $\delta^{13}C$ in bulk and FAME analyses. Thereby indicating the potential for this technique to be in traceability studies.

# Acknowledgements

There are many people that I would like to thank for the support they provided over the last few years, without it, I would not have been able to finish my thesis.

Firstly, I would like to thank the Otago Chocolate Company and Whittakers for providing me with cocoa bean to analysis.

Secondly, I would like to thank my supervisor Professor Russell Frew for his guidance and support he has offered. Along with the time he dedicated to reading and providing feedback on my thesis.

I would also like to thank Dr Robert Van Hall, Dianne Clarke and Dr Amandine Sabadel for all the help they provided in the laboratory.

I would like to thank Dr Kiri McComb for all the advice he gave me, especially regarding statistics.

I would like to thank my research group Dan, Kavindra, Kaushi, Melenaite and Zeki, who were always up for a discussion and a laugh.

Lastly, I would like to thank all my family and friends that have believed in me and encouraged me along the way.

# Table of Contents

# List of Abbreviations

| | |
|---|---|
| C16 | Palmitic Acid |
| C18 | Stearic Acid |
| C18-1 | Oleic Acid |
| C18-2 | Linoleic Acid |
| C19 | Nonadecanoic acid |
| C9 | Nonanoic acid |
| CAM | Crassulacean Acid Metabolism |
| EA | Elemental Analyser |
| FA | Fatty Acids |
| FAME | Fatty Acid Methyl Ester |
| FC | Faraday Cups |
| FID | Flame Ionization Detector |
| GC | Gas Chromatography |
| GC-IRMS | Gas Chromatography coupled to Isotope Ratio Mass Spectrometry |
| IRMS | Isotope Ratio Mass Spectrometry |
| OCHO | Otago Chocolate Company |
| PC | Principal Component |
| PCA | Principal Component Analysis |
| PNG | Papua New Guinea |
| TC/EA | High Temperature Conversion Elemental Analyser |

# List of Figures

# List of Tables

# Chapter 1 – Introduction

## 1.1 Chocolate and Cocoa Beans

**Chocolate**: The definition of what good quality chocolate is, has been a topic of great debate in Europe, so much so that it was even dubbed the Europe Chocolate War (Cidell & Alberts, 2006). The chocolate that is eaten today is very different from the original drink made by the Aztecs and Mayas of Central America. This chocolate drink made from cocoa beans and water was first introduced to Spain in 1520, where sugar was added, and the drink was heated. It was then later introduced to the rest of Europe where it become very popular amongst the rich. However, it was not until the late 18$^{th}$ early 19$^{th}$ centuries that major improvements were made to the drink. The invention and use of steam engines allowed for mass production, which meant chocolate became more affordable and no longer only for the rich.

Cocoa beans are made up of about 50% cocoa solid and 50% fat, known as cocoa butter. The high levels of cocoa butter in the cocoa bean made this drink very fatty and also looked very unappealing due to the melted fat floating on top of the water. In 1828 a Dutch chemist van Houten found a way to remove some of this fat from the cocoa solid by using his invention of the hydraulic press. The cocoa solids could then be ground into a powder improving the consistency of the drink. Finding a use for this extra cocoa butter is what finally lead to the invention of the solid chocolate bar, invented in 1847 by Joseph Fry of England. Fry found that the addition of the left-over cocoa butter to cocoa nibs (deshelled cocoa bean) and sugar produced the right consistency for a solid form of edible chocolate. Milk chocolate was later made in Switzerland by Daniel Peter (Beckett, 2000; Cidell & Alberts, 2006).

Another development from Switzerland that improved the chocolate bar is the process called conching. This process invented by Rodolphe Lindt involves the warmed liquid chocolate being mixed backwards and forwards in a granite trough by a roller often for days. In fact, in Switzerland by law, the chocolate has to be conched for at least 72 hours (Cidell & Alberts, 2006). The outcome of this conching method is chocolate with smoother texture and less bitter taste. It is a combination of these many different processes that make the chocolate we know and love today. However, what people consider to be a good quality of chocolate within

European countries is influenced by which the processes were developed in said country. The two main issues in the European chocolate war were both to do with the ingredients in the chocolate; the percentage of milk and cocoa solids used and the use of cocoa butter equivalents (Cidell & Alberts, 2006).

The use of cocoa butter equivalents, other vegetable fats in replace of some cocoa butter, was debated during the European chocolate war. Chocolate made with 5% cocoa butter equivalents was allowed under the British rules. However, other European countries did not allow any. This caused an issue for British chocolate being sold in other countries as they said that the chocolate was adulterated. There were also issues about the addition of alternative fats enabling the chocolate to be produced more cheaply and therefore making it hard to compete against for smaller chocolatiers. The British ended up winning this argument and were allowed to sell their chocolate with 5% cocoa butter equivalents and up to 20% milk. However, in some countries, the British still need to adjust their advertising to state that it is a chocolate substitute (Cidell & Alberts, 2006).

Today the majority of chocolate is produced by big multinational companies, the three biggest being Mars (USA), Cadbury Schweppes (UK) and Nestlé (Switzerland) (Cidell & Alberts, 2006). However, there are also a growing number of smaller companies that are starting to produce more sustainable labelled chocolate, including Organic, Fair Trade and varietal (cocoa grown in one plantation). In New Zealand, one company that produces varietal chocolate is the Otago Chocolate Company (OCHO) who produce all their chocolate from cooperatives within the Pacific Islands. Each of the OCHO bars is made from single origin cocoa beans (*FAQs*, n.d.)

The simplest form of chocolate is made from only two ingredients; cocoa beans and sugar. Cocoa beans are grown in a particular climate found between 20 degrees north and south of the equator (Beckett, 2000). Countries in this region tend to have high temperatures (~27 °C) and high humidity due to rainfall throughout the year. The evergreen Cocoa tree (*Theobroma cacao*) is native to Central and South America and grows within the shade of the rainforest. When chocolate became popular in Europe, trees were introduced into Africa where they were cultivated in plantations. Currently, cocoa trees are grown in Central and South America, Southeast Asia and West Africa. West Africa is now the largest producer of cocoa beans with

70% being produced here. The biggest producer being the Ivory Coast with Ghana second. Ghana is also known to produce high-quality beans (Beckett, 2000; Diomande et al., 2015).

There are three different varieties of Cocoa trees. Farastero is the main variety used on cocoa farms; it is known to grow fast and have a high yield. The Criollo variety has a much lower yield, but it produces a high-quality bean. The third variety is Trinitario this is a hybrid between the other two. It produces a higher quality bean than Farastero and has a greater yield than Criollo (Beckett, 2000; Diomande et al., 2015). Cocoa beans are grown in pods, each containing around 30-45 beans. These pods are harvested from the trees when ripe and cracked open. The beans are then fermented and dried. Fermentation of cocoa beans tends to be a traditional process that is performed by the farmers to remove the natural pulp on the cocoa beans. It is also a crucial step in the development of the chocolate flavour, it is at this stage where the flavour chemical precursors are formed (Beckett, 2000; Ho et al., 2014).

Once the fermentation and drying of the cocoa beans are complete, the cocoa beans are then sold to companies around the world, where the cocoa beans are used to produce chocolate and other cocoa related products. Chocolate factories all have their own method and recipes for producing their final products; however, within their methods, they all follow the same basic concepts, as shown in Figure 1-1. The first step in making chocolate is the pre-processing of the cocoa bean. Cocoa beans are usually sold as a whole bean, which then needs to be roasted and deshelled. The roasting of a cocoa bean is another important step for flavour development. It is during this stage that the flavour precursors change into the chemicals that give chocolate its taste (Beckett, 2000). The cocoa beans also need to be deshelled as only the cocoa nib is used in chocolate.

The deshelling can occur before or after roasting, depending on the technique used in the factory. Chocoladefabriken Lindt & Sprüngli AG (*Unique Roasting & Grinding*, n.d.) remove the shell before roasting therefore only roast the cocoa nibs where other companies like J.H. Whittaker and Sons, Ltd. (*Beans To Bar*, n.d.) roast the whole cocoa bean including the shell. The process of separating the nib from the shell is called Winnowing. Winnowing usually involves the shell carefully being broken, and air flow is used to remove the lighter shell away from the nib (Beckett, 2000). Once the cocoa beans have been roasted and deshelled, the cocoa nibs are ground down. This process reduces the particle size of the cocoa nib and also releases

the fat from within the cells, forming a cocoa liquor. Once the cocoa nibs have been ground down into the cocoa liquor, additional ingredients such as the sugar and milk powder for milk chocolate can be added. The sugar and milk powder need to be ground down to a small particle size; this can be done either as they are mixed with the cocoa liquor or before. Before the other ingredients are added, the cocoa liquor can be pressed to remove some of the cocoa butter from the cocoa mass (cocoa powder). In the method used by Cadbury's the cocoa liquor is pressed and the cocoa butter is removed. They then make a chocolate crumb with the cocoa mass and add the cocoa butter back at a later date. To make the chocolate crumb, the cocoa mass is mixed with sugar and milk and evaporated (Beckett, 2000; *How Chocolate Is Made*, n.d.). When all the other ingredients have been added and ground down the chocolate liquor is transferred to the Conche.

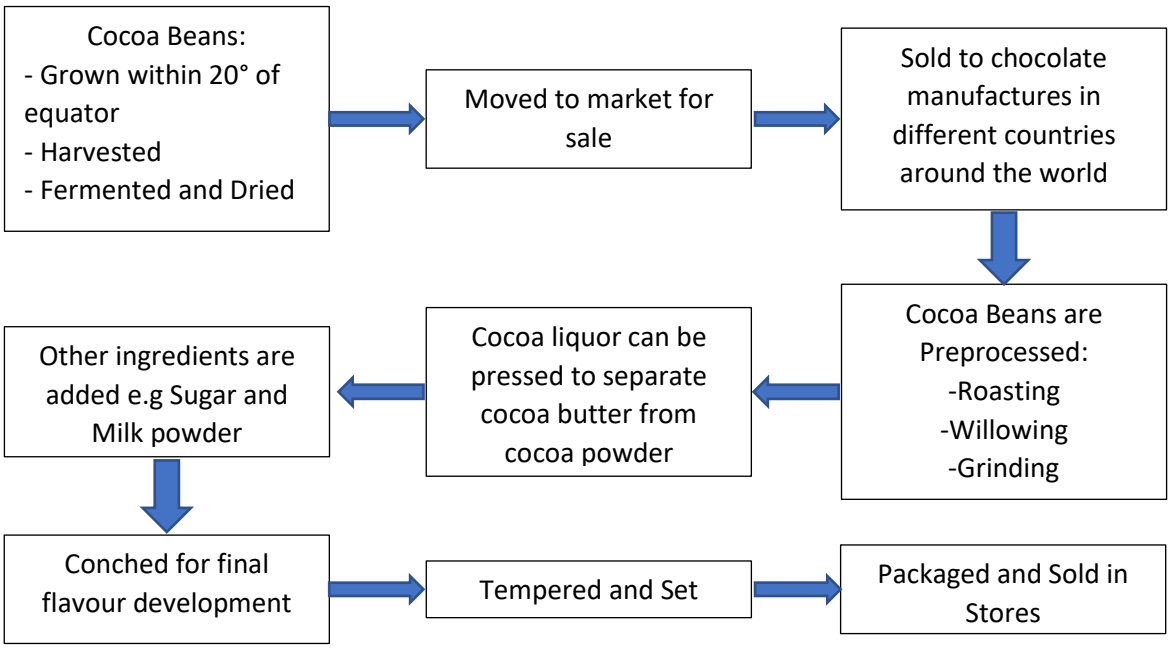


*Figure 1-1. Process of cocoa beans undertake from farm to store*

The Conche, explicitly invented to make chocolate, mixes the warmed chocolate liquor, which helps develop the final flavour by releasing the unwanted acidic flavours developed from the fermentation and roasting processes. It is also a critical step to ensure that the consistency of the chocolate is correct. To help with the viscosity of the chocolate lecithin can be added as an emulsifier and assist in covering the sugar and other particles in fat (Beckett, 2000). The conching process can last for days, as stated earlier in Switzerland, they have to Conche their chocolate for at least 72 hours. The final step in the chocolate making before it is set, packaged and sold is tempering. Tempering is essential to improve the stability and shelf life of the

chocolate. It is a temperature-controlled mixing process of cooling down and reheating the chocolate to ensure that the fat is correctly crystallised. If this process does not take place, then a fat bloom can occur, which is when the chocolate can end up with a white powdery layer. Tempering also gives the chocolate the desired property that allows it to snap (Beckett, 2000; *Finishing with Perfection*, n.d.).

## 1.2 Sustainability

The majority of the worlds cocoa beans are grown on five million smallholder farms in West Africa, South and Central America and some Asian countries (*Food & Farming*, n.d.). Growing the cocoa beans has become a tradition and livelihood for the farmers with the selling of the cocoa beans being their primary or only income. Currently, cocoa beans farms are fighting a losing battle. They are threatened by deforestation, climate change and poverty.

Along with ageing trees which are producing lower yields and also a problem trying to get the younger generation interested in farming ("Everything You Need to Know about Cocoa," 2019; *Rainforest Alliance Certified Cocoa*, 2016). These issues will eventually lead to loss of income for the farmers. One way in which they can improve their situation is to invest in becoming a sustainable farm. Producing products on a certified sustainable farm should lead to higher quality produce as well as higher demand, therefore better returns. It will also allow them to maintain the quality of produce into the future, reducing the risk of loss of income.

Unsustainable farming is a real problem in the cocoa bean farming industry. The industry faces problems such as deforestation, climate change and poverty, which will lead to issues when it comes to the future of the farm. Organisations such as UTZ ("Everything You Need to Know about Cocoa," 2019), Rainforest Alliance (*Rainforest Alliance Certified Cocoa*, 2016) and Fair Trade (*Why Buy Fairtrade Chocolate*, n.d.) have programmes in place to help these farms become sustainable. Some Chocolate manufactures even have their own programme for cocoa farmers to join one example is Cocoa Life which is run by Mondelēz international (*Cocoa Life - Why Cocoa Life?*, n.d.), who are the owners of multiple chocolate brands all of which support Cocoa Life. These programmes all provide many benefits to the farmers and their communities environmentally, socially and economically. The farmers are taught sustainable farming techniques which help them maintain yield and income for the future while protecting the

forests around them. These organisations also work with chocolate manufactures at the other end of the supply chain to ensure that sustainability can be verified from start to finish. Chocolate that has been made with sustainable cocoa beans has a positive impact on the consumers choice, and as the demand for this label grows, the greater the demand for sustainably grown cocoa beans. However, sustainability is not the only problem faced in the cocoa industry; there is also an issue of fraud. With sustainable production sees a higher value in which consumers pay more for. This difference in price can lead to fraud where unsustainably produced cocoa is sold as originating from a sustainable farm to gain the higher price. This leads to the need to be able to authenticate the origin of the cocoa beans.

A farm can become certified as sustainable by following the procedures that the programs put in place. To ensure that the farmers are sticking to the procedures they are audited. The different organisation all have their own criteria but with the same end goal to help the farmers become more sustainable for the future. One of the issues with unsustainable farming is deforestation. As the cocoa trees age, they become less efficient and produce fewer cocoa beans. To increase their yield farmers then cut down protected forests in order to create space to plant new trees. These trees can include hybrid trees which need more sunlight; therefore, fewer shade trees and greater use of pesticides, both of which can impair natural habitats. (Rainforest Alliance Certified Cocoa, 2016; Walz, 2019). To help protect the forests from deforestation farmers that are certified with UTZ, who are part of the rainforest alliance, are asked to replace some of the ageing trees on their farm with new plants. Ageing trees decrease the yield of cocoa beans grown by replacing them with new trees; it helps increase the yield while not cutting down protected rainforests to make new space.

Another issue that cocoa bean farmers can face is climate change. Cocoa beans can only be grown in a narrow climate range. Therefore, the increasing temperatures and unpredictable rainfall associated with climate change can cause an issue with the quantity and quality of the cocoa beans grown over time. As a part of their program, UTZ have encouraged farmers to plant native shade trees on their section. These trees help lower the temperature by providing shade (Walz, 2019). In following these procedures, it doesn't just help keep the environment sustainable, but it helps increase the yield and therefore, income for the farmers, making them also economically sustainable.

Many of these organisations, such as UTZ and fair trade, who aim to help the farmers become certified sustainable, have control over the whole supply chain. They certify the supplies which purchase the cocoa beans to help ensure that the farmers are getting a fair price ("Traceability System," n.d.). Fair-Trade has a set minimum price levels for cocoa bean purchase from the farmers. Also, they add a premium on top to go towards helping the farmers communities with education and healthcare as an example (Why Buy Fairtrade Chocolate, n.d.). When it comes to the community, the main focus is to reduce child labour and promote gender equality. Child labour is a massive problem in the cocoa-growing industry; it coincides with poverty. Children need to work to help out their parents who do not earn enough to support the family. Farmers hire them as they are cheaper labour than adults. Farmers also get their own children to help out, so they do not have to hire anyone else. With children having to work, it means they are not getting an education. Another factor that can lead to children working is that the schools are too far away in the cities and parents are unable to afford to send them. (Diakite, 2017; Fairtrade and Child Labour, n.d.)

To achieve the goal of eliminating child labour, the different organisations help ensure a stable and sustainable income for the farmers and their workers. Thus, reducing or even eliminating the need for children to work to help their families. They also help build schools in the area or ensure there is transport so that the children can go to school and get an education (Diakite, 2017; Fairtrade and Child Labour, n.d.). Another social issue that the organisations focus on is gender equality. The cocoa industry is very much male-run, and these organisations want help to ensure females have equal rights. They help promote gender equality by providing courses for women to learn leadership skills and teach the community of equal rights. They encourage women to help make decisions and support them to own farms (Cocoa Life - Cocoa Life Is Empowering Women, n.d.; Francois, 2019). These organisations help farmers and their communities to ensure they have a more sustainable future environmentally, socially and economically. The chocolate manufactories that buy cocoa from these farms have the knowledge that they have been produced sustainably. Chocolate that has been made from cocoa beans produced sustainably will have the sustainability logo from the organisation that certified the farms. The logo provides information to the consumers about the sustainability of the cocoa beans.

Do sustainable labels have any impact on the consumers choice when it comes to purchasing chocolate? There have been multiple studies conducted to find what impact of these sustainable

labels and consumers choice when it comes to purchasing chocolate. However, one of the issues raised was that participants did not always know what the labels were or what they meant. One study by Rousseau (2015) asked their participants to identify three different sustainable labels and found that the majority could only identify the fair-trade logo. They also found that about half of their participants did not believe in the creditability of the labels. In a study by Silva et al. (2017), the participants also lacked knowledge of the sustainability labels found on chocolate. This lack of education and knowledge of the labels indicates that before the participants began the study, there was little influence on their choice of chocolate. These two studies show that there is also a need for more education about sustainability. Providing information about the benefits to the cocoa bean farmers, their community and the environment. Along with certified labels from the sustainability organisations chocolate also can have labels signifying if it is organic and the origin of the cocoa beans. All these sustainability labels can have an impact on the consumers choice when it comes to purchasing the product. For the labels to impact a consumer's choice, the consumer needs to have confidence in them. As stated earlier, the participants in Rousseau's study (2015) did not believe in the creditability of the labels. Therefore, these sustainability issues need to have a measurable way of verification and enforcement. A lesson learnt from the horsemeat crisis of 2013, and the Elliot report is that this needs to be supported at all levels of the food industry and government (Elliott, 2014).

Young and McCoy (2016) published a two-part study by focusing on what influenced the chocolate purchases of millennial Midwestern Americans. They found that in their focus group section of the experiment, for the younger age group (18-25 years) taste was the leading influencer. Whereas the older age group (26-35 years) showed a greater preference for the social factors (organic, ethically sourced). However, most still did not buy the chocolate which displayed positive social factors. Out of those who did it was mostly a guilt-reducing purchase. The second part of the experiment was a choice experiment; it confirmed that what was in the chocolate (ingredients, fat and sugar content) that influences the taste, was a more significant influencer to the choice of chocolate they purchased over social factors. However, there was a small group that found social factors more important.

A similar study conducted by Silva et al. (2017) on the impact of a sustainably-produced label on purchasing chocolate for Brazilians had similar results. The experiment consisted of a blind sensory analysis, then a second analysis with sustainability labels provided. They found that the sensory test had the most significant impact on the intent to purchase. However, there was a

positive impact on the willingness to purchase shown for products showing the sustainability labels. As the intention to buy the product was increased once the sustainability labels were provided. Rousseau (2015) conducted a study to see the impact of sustainability labels on chocolate in Belgium. Their results also showed that when it came to what chocolate the participants were going to buy taste was the most important factor. However, the fair-trade label did have a positive impact on the participant's willingness to pay more for the product. Unlike the Fair-trade labels, it was found that the products labelled as organic did not produce the same positive impact on the participant's willingness to pay more. In fact, for some participants, it even harmed their willingness to pay more. From these study's it is clear that from a consumer's point of view, the taste is the most important factor they consider when deciding what chocolate to buy. However, there is still a positive impact and even greater willingness to pay more for products that have been produced sustainably.

There are many benefits for farmers to start to grow their products sustainably, including positive impacts on consumers choice, leading to a higher demand which can increase the price. When the purchasing choice is made on ethical rather than inherent (e.g. taste) grounds, it becomes crucial for consumers to be confident in the authenticity of the cocoa beans because they are relying on information they cannot verify. The sustainability programmes have traceability procedures to trace the origin of their products so it would stand to reason to assume that this was nothing to worry about; however, this may not necessarily be the case.

There are multiple steps in the supply chain for the cocoa beans from the farm to the final chocolate product, as seen earlier in Figure 1-1. With there being multiple steps, this provides an opportunity for fraud once the cocoa beans have left the farm. There is also a lack of law enforcement to stop this from happening. Amankwah-Amoah et al. (2018) highlight some of the issues within the supply chain in Ghana. They found that there was a big problem with smuggling out of the country. Uncontrolled transhipment can lead to fraud by stating the beans come from a different country from which they were grown. The smuggling was found to be conducted by farmers who found it cheaper and easier to smuggle the beans across to Ivory coast than to move them to their closest market. Smuggling was also being conducted by thieves who target all sections of the supply chain. There is very little stopping this smuggling from occurring as the law enforcement accept bribes to turn a blind eye to the issue. Amankwah-Amoah et al. (2018) also found that there was a problem with the weighing scales being adjusted so that the bags would end up underweight. To bulk up their bags, farmers would mix the quality

of the bags by having the poor-quality beans at the bottom of the bag and high quality at the top to get the higher price. Cocoa beans from Ghana sell at a premium (Amankwah-Amoah et al., 2018) which means that with the knowledge of some issues with the industry in Ghana it is important to be able to track the authentication of the origin of cocoa beans.

## 1.3 Authenticity

Confidence in the authenticity of food is becoming increasingly important to consumers, and this has led to multiple studies on how to authenticate food products. As stated previously, the main concern around the authenticity of cocoa beans is to ensure that they are being grown sustainably. A complication in how to determine the origin of the cocoa beans is that cocoa beans are usually sold as a manufactured product, chocolate. Mixing complicates the application of many authentication methods as chocolate is a combination of ingredients with dark chocolate being the simplest made with only cocoa beans and sugar. Other varieties of chocolate become much more complex with ingredients including the addition of milk for milk chocolates. There have currently been a few studies on determining the origin of cocoa beans. Most of which only focus on the cocoa beans themselves, however, some also include dark chocolate samples.

Some previous studies have looked at techniques that relate to the metabolism of the cocoa bean. These include using $^{1}$H NMR to study the metabolic profile (Caligiani et al., 2014); the fermentation level was found to be the primary influencer. There was some relation to the origin of well-fermented beans. However, this was found to be based on the variety of bean grown in different regions. Jinap & Dimick (1990) investigated the acidic characterisations, including pH, titratable acidity and concentration of volatile and non-volatile acids. They found samples were able to be classed into three groups Brazil and Far eastern countries, Central and South American, and West African. A study by Chaiseri & Dimick (1989) looked at the physical characteristic of hardness in cocoa butters along with the triacylglycerol compositions. Their results showed that the triacylglycerol compositions which affect the hardness of the cocoa butters were related to the climate which allowed for three groups to form South American, Asia and Oceania and North American and African cocoa butters.

Other studies that have been conducted on cocoa beans and chocolate products have focused on the chemical components within the cocoa beans, such as fatty acid profile, multi-elemental profiling, volatile components and isotope ratios. They have then applied multi-variate statistical methods to the data to help classify it based on origin. In a study by Torres-Moreno et al. (2015) the nutritional composition and fatty acid profile of cocoa bean and chocolate samples from Ghana and Ecuador were compared to determine if there was a difference between the samples of different origin. The analysis of the nutritional composition showed that there was a difference in fat, carbohydrate, fibre and moisture composition in cocoa beans from Ecuador to those from Ghana. This difference in nutritional composition in the cocoa beans was also present in the chocolate with chocolate from Ecuadorian cocoa beans having higher fat content than chocolate made from Ghanaian cocoa beans. These differences were due to the origin and not the processing methods of the chocolate. In the fatty acid profile of the cocoa beans C16, C18 and C18-1 were found to be the most quantitively important fatty acids with their concentration around 25%, 33% and 34% in the cocoa bean respectively. Torres-Moreno et al. (2015) found from the fatty acid profile that seven of the 15 fatty acids in the cocoa beans showed a significant difference due to the origin of the cocoa bean. C16 and C18 were both included in these seven fatty acids. Even with the differences in seven of the fatty acids, it was noticed that the fatty acid profiles for all of the cocoa beans were very similar. The chocolate samples were found to only have four fatty acids with a significant difference between the origins. These differences were in the four most quantitative fatty acids C16, C18, C18-1 and C18-2.

Cambrai et al. (2010) analysed the volatile compounds in chocolate to determine if it was possible to use this technique as a way to determine the geographical origin of the cocoa beans used to make the chocolate. The samples used in this study were from 8 different countries. The results from the Principal Component Analysis (PCA) (Figure 1-2) showed that it was possible to completely distinguish chocolate made from Caribbean cocoa from all the other origins. The chocolate samples from Madagascar and Western Africa cocoa formed two clusters from the PCA. However, there was some cross over between these two clusters and the mixed zone. The mixed zone cluster contained samples from Southern America and Madagascar, along with samples from other countries. The results of the Factorial Discriminant Analysis (FDA) also distinguished the Caribbean group from all the other groups. The others were grouped in Africa, South America and Madagascar, these each formed a cluster. Cambrai et al. (2010) performed a blind test with new samples to test the clusters and found >90% probability for the sample to

belong to a group. However, >75% of the results were a true positive identification to the original group. Two samples produced a false positive as they were assigned to a group that they did not correspond too. Some results produced a false negative due to little separation between some groups. In this experiment, the only country that had its own group was Madagascar all other groups contained samples from multi countries in the same region.

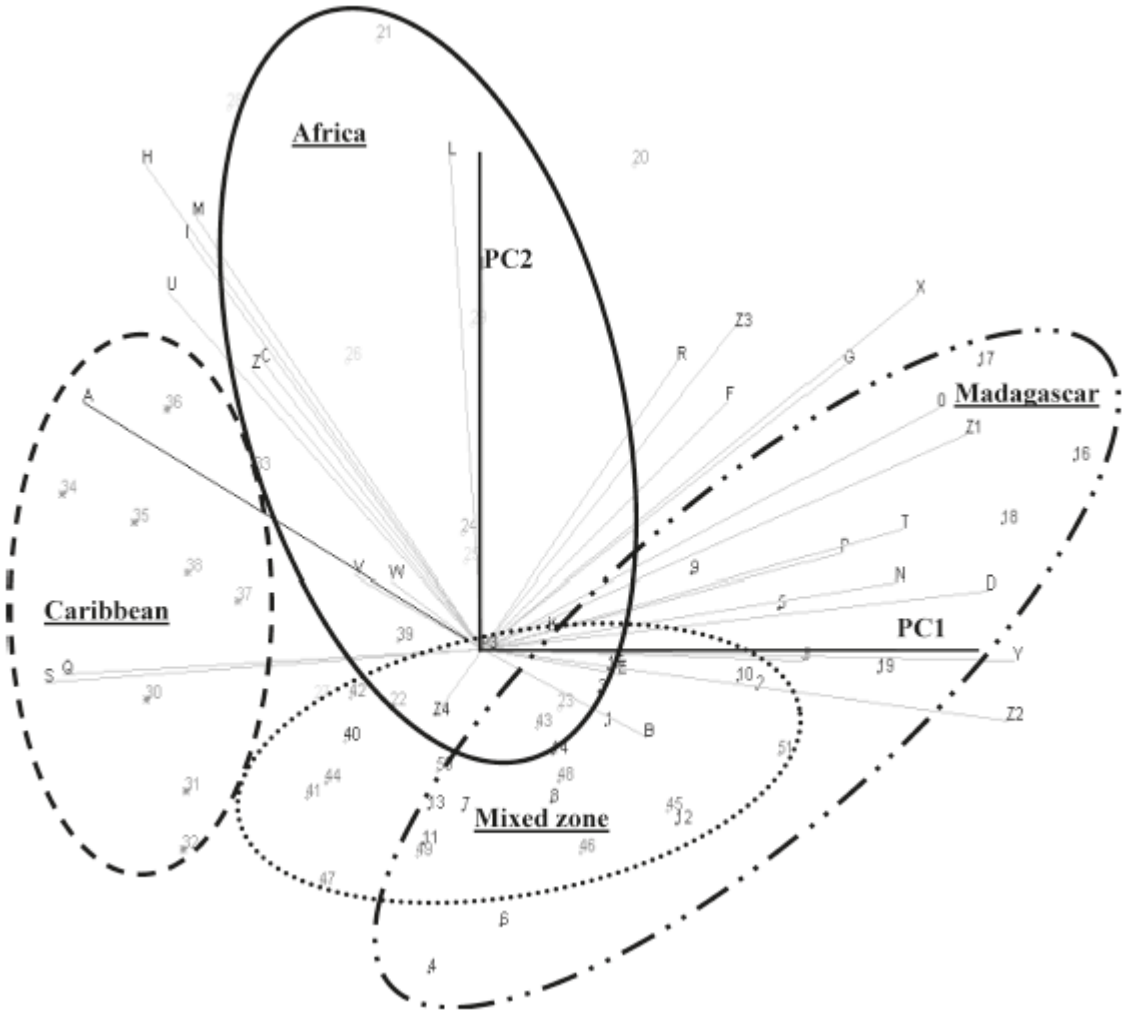


*Figure 1-2. Plotted PCA numbers 1-51 represent chocolate samples, vectors A-Z4 represent chromatographic peaks. Figure taken from Cambrai et al (2010)*

Bertoldi et al. (2016) investigated the geographical traceability of cocoa beans using a multi-elemental fingerprinting approach. They analysed samples from 23 different countries and created a model with 29 elements and found it possible to discriminate the samples by five subcontinental origins, West and East Africa, Asia, and Central and South America. A leave one out test was conducted on samples from countries with more than two samples. The results from this test showed that 93% of samples were correctly classified to the subcontinental origin.

This study then applied the same technique to dark chocolate to see if they could classify the chocolate into the same origins. With chocolate being packaged for commercial use, there is the possibility of contamination. There was a correlation seen between the elements in chocolate to that in the cocoa bean. To account for the sugar component in the chocolate, results were recalculated based on the percentage of cocoa solids as stated on the package. They applied the model used for the cocoa beans for geographical origin to the dark chocolate. The results showed that for both the cocoa beans and chocolate together, 96% of the samples were classified correctly. Therefore, multi-elemental fingerprinting does provide discrimination of origin for both cocoa beans and dark chocolate. These studies all prove that it can be possible to determine the origin of cocoa beans from the finished chocolate product.

All these studies show that they were able to see some differentiation between origins. However, both the studies from Bertoldi et al. (2016) and Cambrai et al. (2010) were unable to separate the origins by country rather just the regions. This is an issue when it comes to trying to prove authentication for samples from countries within the same region. The method of multi-elemental fingerprinting used by Bertoldi et al. (2016) did show success in the separation of origins; however, it does have some limitations. With cocoa beans being sold commercially as a component in the manufactured product of chocolate, it is exposed to contamination during this process. The contamination can arise from the instruments used to turn the cocoa beans into chocolate as well as the packaging. Chocolate is manufactured from multiple ingredients the products in the study were dark chocolate which is the simplest form of chocolate containing only cocoa beans and sugar. However, chocolate can be made with other ingredients added to it, such as milk for milk chocolate or other nuts and flavours. For these products, it would be much harder to determine where the cocoa beans came from using multi-elemental fingerprinting.

The study by Torres-Moreno et al. (2015) is limited in what it shows due to only comparing cocoa and chocolate samples from two different origins. Even though there were significant differences in the fatty acid profiles between these two origins when comparing profiles from other countries, overlap might occur. There are only three dominant fatty acids in cocoa beans, and only two showed a significant difference between the two countries, therefore to add more variables to the analysis relies on fatty acids of naturally low concentrations within the cocoa bean. This would become an important issue for using this technique as a way of discrimination between origins. When variables are at a low concentration, it becomes harder to develop the

correct quality control factors that are crucial for comparing samples between datasets. When the fatty acids of the chocolate samples were compared, it was seen that only the four most quantitative fatty acids had any significant difference between them. This is different from what was found in the cocoa beans, which would complicate determining the origin of cocoa in chocolate from a database based on cocoa bean values. Fatty acid profiles for origin authenticity have a significant limitation when it comes to analysing any chocolate that is not dark as the milk in milk chocolate also has fatty acids which would then complicate the fatty acid profile. All three of these studies on the origin of cocoa in chocolate show that there are differences between samples; however, these differences tend to be in regions of countries, not just individual countries. They are also limited in that they have only studied the simplest form of chocolate; dark, and if other ingredients were added to the chocolate, there would be an issue in separating the results of the cocoa beans in chocolate with the other ingredients.

Not only is there an issue with there being additional ingredients to cocoa beans and sugar added to make chocolate, but some manufactures also add other fats along with cocoa butter to the chocolate. Spangenberg and Dionisi (2001) conducted a study comparing cocoa butter from cocoa beans with cocoa butter equivalents. When chocolate is made, it can sometimes have additional products added, which include the cocoa butter equivalents. These are fats which have similar characteristics as cocoa butter (Spangenberg & Dionisi, 2001). The study looked at the stable carbon isotope and composition of fatty acids in cocoa butter, cocoa butter equivalents and mixtures to see if it was possible to identify the cocoa butter equivalents within a chocolate mixture. Both the bulk isotope ratio and the individual isotope ratio of the fatty acids within the cocoa butter and cocoa butter equivalents were analysed. The results showed that it was possible to separate the cocoa butter from the cocoa butter equivalents. Spangenberg and Dionisi (2001) then created a model which included the fatty acid concentrations along with the bulk and compound-specific isotope ratios to determine if it would be possible to detect cocoa butter equivalents when they are added to cocoa butter in low concentrations. Mixtures of 15% cocoa butter equivalents in cocoa butter were theoretically calculated to determine if the mixtures could be separated from the pure cocoa butter. The results showed that most mixtures made with 15% cocoa butter equivalents could be separated out from the pure cocoa butter. This study shows that if cocoa butter equivalents have been added to the cocoa butter at amounts as low as 15%, this can be detected. This is important for chocolate samples in which manufacture may produce chocolate with cocoa butter equivalent which are often a cheaper option.

Isotope ratio analysis has also been explored to see if it can be used as a way to determine the origin of cocoa beans. This technique was analysed in a study by Perini et al. (2016). In the study Perini et al. analysed the isotope ratios of C N H O and S in cocoa beans, to determine if the isotope profile could be used to trace the cocoa beans back to their origin. The results from the box and whisker plots Figure 1-3 showed that when looking at the isotope ratios of the different origins individually, it was not possible to separate all the different locations. However, there were some differences seen between some of the countries. For the $\delta^{13}C$ values, there was some discrimination between samples from the Northern Hemisphere and those from the Southern Hemisphere. There were also some differences seen from within a region as it was possible to discriminate Tanzanian from Ugandan samples, both of which are in East Africa.

The analysis of the $\delta^{2}H$ and $\delta^{18}O$ showed that there was a correlation between the $\delta^{2}H$ and $\delta^{18}O$ in the cocoa beans. Where $\delta^{2}H$ and $\delta^{18}O$ are the isotope ratios compared to a standard for H and O respectively. The isotope values were also correlated with rainfall perception from the Global Network for Isotopes in Precipitation database. When comparing the $\delta^{2}H$ and $\delta^{18}O$ for the samples (Figure 1-3), it was seen that Uganda was significantly different to all other countries with a very high value for both $\delta^{2}H$ and $\delta^{18}O$. While Peru, Indonesia and Papua New Guinea all had relatively low values. The other countries in the study all had similar values in between. The $\delta^{15}N$ and $\delta^{34}S$ also showed some variation between the countries. It was found that the West African countries, along with the Dominican Republic, had high $\delta^{34}S$ values compared to all the other countries. Within Central America, there was a significant difference between the Dominican Republic and Mexico. $\delta^{15}N$ also showed variations between countries of the same region. In West Africa, Ghana and Ivory Coast had higher values compared with Nigeria, São Tomé and Principe and Sierra Leone. Individually each isotope ratio was unable to completely separate the cocoa beans based on their origin apart from Uganda being separated from all other countries for $\delta^{2}H$ and $\delta^{18}O$. Like some of the other techniques for determining the origin, it was seen that regions could be separated from others. However, overall, the individual isotope ratios are not able to determine the origin of cocoa beans.

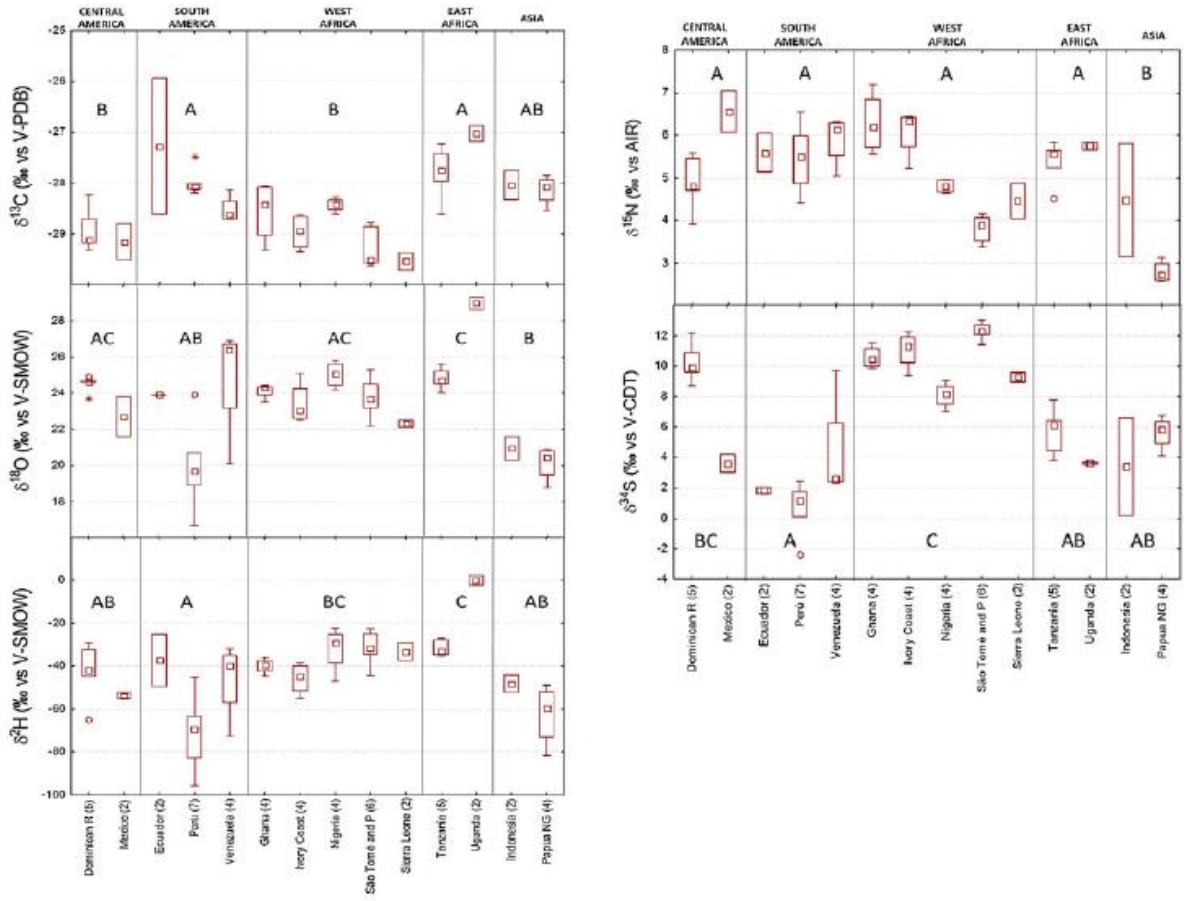


*Figure 1-3. Box and Whisker plots for each isotope ratio ($\delta^2H$, $\delta^{13}C$, $\delta^{18}O$, $\delta^{14}N$ and $\delta^{34}S$) measured on cocoa beans from different countries. Figure taken from Perini et al (2016)*

Perini et al. (2016), conducted a multi-variate canonical discriminant analysis to determine if there is more significant discrimination between the cocoa samples when the isotope ratios values are combined. The statistical model, which had been applied to all isotope ratios, was able to discriminate samples based on their isotope values, especially loading on $\delta^{15}N$, $\delta^{13}C$ and $\delta^{34}S$. The Papua New Guinea samples were able to be discriminated from all other samples based on their use of synthetic fertilization. Sao Tãmé and Peru also had high discrimination from the other countries. Isotope ratios as a technique have been shown to have the ability to be able to discriminate some countries of origin from others. The sample size used by Perini et al. (2016) was small, with samples from some countries only having two cocoa beans. Conducting this experiment with more samples would help to determine if more countries could be discriminated out using isotope ratios.

A study by Diomande et al. (2015) also used isotope ratios to classify cocoa beans based on origin. In this study, Diomande et al. measured the $\delta^{13}C$ and $\delta^{15}N$ as well as the percentage of

N and C on the cocoa bean from 24 different geographical origins. The whole cocoa bean was analysed as well as sub-matrices of the cocoa bean. These sub-matrices included the Shell and Cotyledons, which was then broken down to lipids and cocoa powder (defatted cocoa bean). The cocoa powder was further broken down to theobromine and proteins. All isotope ratios were measured using an Elemental Analyser coupled to an IRMS. A PCA was conducted on the data and found that the first two PC (Principle Components) only explained 46% of the variability. A sparse PCA was then conducted this limited the number of variables to the most discriminating variable of the original first two PC. The reduced variables included $\delta^{13}C$ and $\delta^{15}N$ of proteins, cotyledons, defatted cocoa, shells and $\delta^{13}C$ of theobromine and $\delta^{15}N$ of whole cocoa beans. They found that two clusters formed one with samples from Africa and the second with samples from other countries. The $\delta^{15}N$ values of these variables showed some discrimination based on origin, as seen in Figure 1-4. It is seen that the samples from Africa have the highest values compared with the other origins. With the majority of the samples in the study being from the Ivory Coast in Africa, an analysis of three different towns within the Ivory Coast was conducted. This analysis consisted of conducting a supervised Partial Least Squared Discriminant Analysis. The results showed that the $\delta^{15}N$ of theobromine and defatted Cocoa along with the $\delta^{13}C$ of proteins, defatted cocoa and lipids were able to discriminate the cocoa beans of these three towns from each other.

The studies by Perini et al. (2016) and Diomande et al. (2015) have shown that isotope ratios can be used to discriminate cocoa beans based on origin. The isotope technique used in these studies measured the ratio of the bulk samples, which means if this method was used to analyse chocolate, then the sugar and other ingredients would also contribute to the isotope ratio. However, the study by Diomande et al. went on to measure the isotope ratios of sub-matrices within the cocoa beans which were also seen discriminate based on origin. Which could lead to analysing the isotope ratio of these sub-matrices of cocoa beans within a chocolate product.

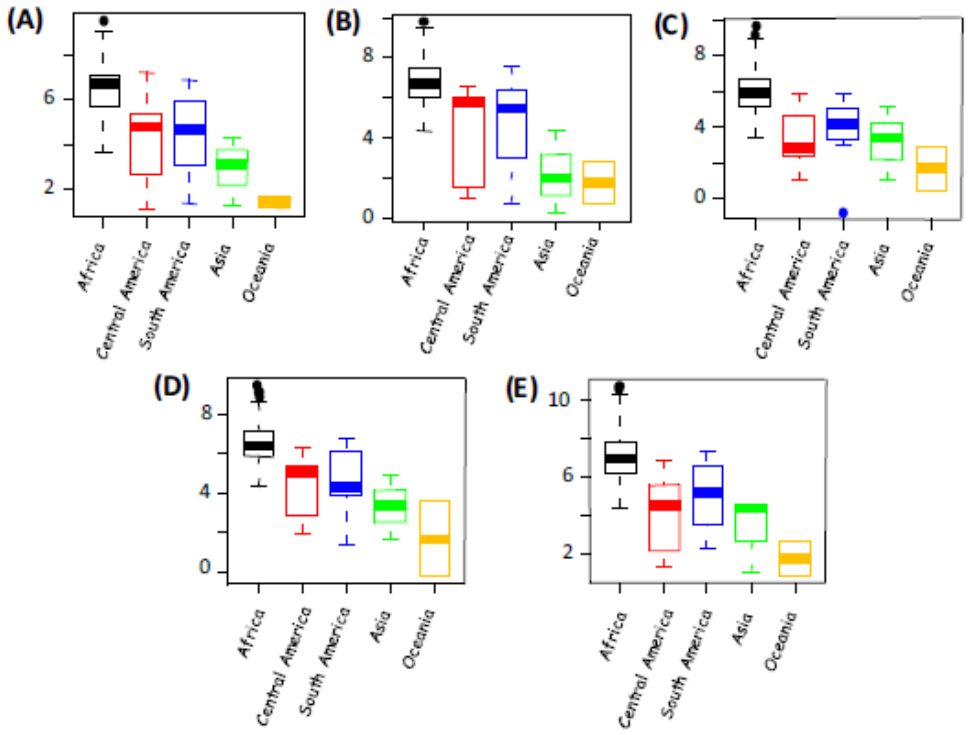


*Figure 1-4. Boxplots of $\delta^{15}N$ (A) Cocoa beans, (B) Cotyledons, (C) Shells, (D) defatted cocoa, and (E) Protein for Africa, Central America, South America, Asia and Oceania. Figure from Diomande et al., 2015.*

## 1.4 Composition of Cocoa beans

Cocoa beans are made up of around 50% cocoa butter and 50% cocoa solids. Torres-Moreno et al. (2015) found that the major nutrients in cocoa beans were fat (>40%), carbohydrate (>32%) and proteins (12-13%). There are four main fatty acids shown in Table 1-1 that make up 98% of the total fatty acid amount in the cocoa butter. These fatty acids form the three main triglycerides in cocoa butter. Glycerol-1,3-dipalmitate-2-oleate (POP), glycerol-1-palmitate-2-oleate-3-stearate (POS) and glycerol-1,3-distearate-2-oleate (SOS) (Jahurul et al., 2013).

*Table 1-1. Major fatty acids in cocoa beans. Table adapted from Jahurul et al. (2013)*

| **Major Fatty acids in cocoa beans** | **Formula** | **Percentage of fatty acid** |
|---|---|---|
| Palmitic acid | $C_{16}H_{32}O_2$ | 25 - 33.7 % |
| Stearic acid | $C_{18}H_{36}O_2$ | 33.7 - 40.2 % |
| Oleic acid | $C_{18}H_{34}O_2$ | 26.3 - 35 % |
| Linoleic acid | $C_{18}H_{32}O_2$ | - 3 % |

## 1.5 Fatty acid – isotope ratios

Compound specific isotope ratio mass spectrometry provides the ability to measure isotope ratios of individual compounds within a substance. One class of compounds that has previously been analysed this way from food sources are the fatty acids. By extracting out the fatty acids from the source and derivatising them to form an ester, allows the individual fatty acids to be separated through the GC column. The fatty acid esters are then converted to the required gas ($H_2$ or $CO_2$) in a reactor before entering the isotope ratio mass spectrometer. Which then provides the isotope ratio of the target element for each fatty acid. This technique has been explored in previous studies as a way of traceability for different food sources.

Multiple studies have been conducted using fatty acid isotopes ratios on food oil, focusing on the origin of the oil. In a study by Faberi et al. (2014), fatty acid $\delta^{13}C$ along with bulk $\delta^{13}C$ and fatty acid composition were used to determine the origin of Olive oil from different regions of Italy. The samples were divided into three different zones based on pedoclimatic factors but for the most part, Southern, Central and Northern Italy. Both PCA and PLS-DA (partial least squares discriminant analysis) were used to analyse the data. The results showed that when the fatty acid composition was removed, greater separation of the three groups was achieved. This shows that the combination of bulk and fatty acid provides a good base for the geographical origin of olive oil samples from Italy. Another study that analysed the origin of Olive oil was by Bontempo et al., (2019). In the study, Bontempo et al. analysed the $\delta^{13}C$ and $\delta^{2}H$ on the individual fatty acids of olive oil samples from around the world along with the bulk isotope data. The study aimed to determine if it was possible to separate samples from European countries with those from non-European countries. The result from a PCA showed that although samples could not be separated based on European and non-European countries, samples did tend to cluster according to the country of origin. A Random forest classification was also run, and this analysis showed that when the fatty acid isotope data was added to the bulk data, there was a significant improvement in the classification abilities. Another oil that has been analysed is Camelina sativa oil in a study by Hrastar et al., (2009). The study compared samples from North America with samples from Central and Northern Europe. Using the $\delta^{13}C$ of four fatty acids and bulk oil, the samples from North America could clearly be distinguished from the samples from Europe. They were, however, unable to sub-cluster the samples from Europe. All of these studies have shown that the analysis of fatty acid isotope ratios can provide information on the origin of oil with significant improvement from analysing the bulk isotope data alone.

The fatty acid isotope ratios have also been used for animal base products with Ehtesham et al., (2013) applying the technique to bovine milk. Ehtesham et al. analysed the isotope ratios of New Zealand milk powder using both bulk and individual fatty acids isotope ratios to see if there was a correlation with the $\delta^2H$ values of precipitation in the New Zealand climate. The bulk ratios were found to correlate with the precipitated water values and that three of the fatty acids also correlated with the precipitated water values. Ehtesham et al. also investigated whether the milk powder could be discriminated by origin from their bulk and fatty acid values. The $\delta^2H$ of the bulk and fatty acids Oleic, Palmitic, Butyric and Myristic were able to separate the samples between the North and South Islands of New Zealand along with distinct clustering within the North Island samples.

Some studies have shown some success when analysing isotope ratios of fatty acids as a way to determine the origin of seafood products from off the coast of China. In a study by Liu et al. (2017), the $\delta^{13}C$ isotope ratios of fatty acids were analysed for the sea cucumber (A.*Joponicus*). Samples from multiple locations off the coast of China were analysed over two seasons spring and autumn. From this study, it was shown that the majority of samples were able to be separated based on origin using both PCA and canonical discriminant analysis. Except for two origins which were found to have similar values for the samples collected in spring. These two origins were close to one another and likely provided similar food sources. A study Zhang et al., (2019) showed similar results for the origin of scallops. The $\delta^{13}C$ of individual fatty acids for three species of scallops from the coast of China were analysed. The $\delta^{13}C$ fatty acid results were able to distinguish most origins from each other with the use of PCA. With the exception of three origin pairings across both season and all three species analysed, where some overlap was seen. The results of a linear discriminant analysis used to classify the origin of samples correctly had an average accuracy of 85.3% for the three species. The study also found that combining the $\delta^{13}C$ fatty acids results with the fatty acid profiles, which also could not distinguish between all origins, that the combine results were able to distinguish between all origins using linear discriminant analysis (LDA) with a 100% correct classification rate.

These studies have shown some success in using the isotope ratio of individual fatty acids as a way to discriminate different foods sources based on origin. Stable isotope ratios are a great technique to use for differentiating samples based on origin for a pure product, more information on how stable isotopes can be used for discriminating between origin is found in chapter 2. They are not, however, very effective when it comes to mixtures. This is where the

ability of compound-specific IRMS comes into effect. This technique allows targeted analysis. In the case of cocoa beans, which are often sold as a mixed product such as chocolate with additional ingredients, sugar as a minimum. Analysing the FA of the cocoa beans from within the chocolate could provide the ability to determine the origin of cocoa beans within a chocolate product. This current study will look at applying this technique to cocoa beans as a means of determining if the FA isotope ratios can provide origin information. Then to chocolate samples to see if it is possible to use the FA from the chocolate to determine the cocoa beans origin.

## 1.6 Aim of the current study

Aim: Investigate if stable isotope measurements can be used to verify the origin of cocoa extracted from chocolate.

The previous studies on authentication of cocoa beans have shown that different techniques do have some discriminating ability when it comes to determining the origin of cocoa. They also show some promise with the ability to apply these techniques to chocolate in its simplest form. Both fatty acid profiling and stable isotope ratios, specifically carbon and hydrogen, have shown to be useful techniques for this work. The current study will look at combining these two techniques to see if it is possible to determine the use of isotope ratio analysis on the fatty acids within cocoa beans. Then applying this technique to chocolate samples to determine the origin of the cocoa beans used.

This study will be conducted in three parts. The first will be to measure the bulk $\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$ of the cocoa beans. These results will be used to see if it is possible to discriminate the cocoa samples based on country of origin. The results will also help give an understanding of the variability within and between the countries.

The second part is to extract the fatty acids of the cocoa beans and measure the $\delta^{2}H$ and $\delta^{13}C$ of these fatty acids using gas chromatography isotope ratio mass spectrometry (GC-IRMS). This process involves the formation of fatty acid methyl esters (FAMEs), which are separated through the GC column before entering the IRMS. This technique allows for the isotope ratios of the fatty acids to be measured individually. This study will focus on the three most dominant

fatty acids found in cocoa beans Palmitic acid (C16), Stearic acid (C18) and Oleic acid (C18-1).

The final part is to evaluate the isotope ratio of cocoa beans within a chocolate sample. This will be achieved by extracting fatty acids from chocolate and measuring the $\delta^{2}H$ and $\delta^{13}C$ of these fatty acids. The data will be compared to the results of the cocoa bean fatty acids isotope ratios.

The objectives of this study are:

1. Determine if stable isotope ratios do reflect the geographical origin of cocoa beans by:
    a. Acquiring samples of cocoa beans from different geographical origins
    b. Prepare the beans and measure the H, C, and N isotope ratios on them
    c. Explore the data to determine if the results cluster according to origin classification

2. Evaluate if stable isotopes of fatty acids from beans reflect origin.
    a. Extract the fatty acids from the beans and purify
    b. Derivatise the fatty acids to form their respective methyl esters
    c. Measure the H and C isotope ratios on the individual FAMEs
    d. Explore the data to determine if the results cluster according to origin classification

3. Extract fatty acids from chocolate and evaluate if they reflect the origin of the cocoa beans.
    a. Make chocolate with beans from a known origin
    b. Extract and derivatise the fatty acids from the chocolate
    c. Measure the H and C isotope ratios on the individual FAMEs
    d. Explore the data to determine if the results cluster according to origin classification
    e. Test for the influence of lecithin

# Chapter 2 – Isotopes and IRMS

## 2.1 Stable Isotopes

All atoms are built up of three different components; protons, electrons and neutrons, except for Protium which contains zero neutrons. An element is defined by the number of protons it contains. The atomic mass of an element is calculated by the sum of the number of protons and neutrons. Isotopes are different forms of an element that differ only in the number of neutrons hence a difference in the atomic mass. The numbers of protons and electrons between the different isotopes of an element remain the same. Chemical properties, such as electronegativity and number of covalent bonds formed, are reliant on the numbers of protons and electrons. Therefore, these properties are consistent between the isotopes of a given element. There are two classes of isotopes unstable and stable. Unstable isotopes undergo radioactive decay; this process causes variability in the abundance of the isotopes over time. For stable isotopes which have not been seen to undergo radioactivity, the relative abundance is fixed globally, with usually a much greater abundance of the lighter isotope. Table 2-1 presents the stable isotope abundance for H, C, N and O. However, even though the relative abundance is fixed for each isotope, it is not equally distributed (Criss, 1999; Sharp, 2007; Wakefield, 2013).

*Table 2-1. The global abundance for the stable isotopes of Hydrogen, Carbon, Nitrogen and Oxygen. Adapted from Criss 1999*

| Element | Isotope | Abundance (atom %) |
|---|---|---|
| **Hydrogen** | $^{1}$H (Protium) | 99.985 |
| | $^{2}$H (D or Deuterium) | 0.015 |
| **Carbon** | $^{12}$C | 98.90 |
| | $^{13}$C | 1.10 |
| **Nitrogen** | $^{14}$N | 99.63 |
| | $^{15}$N | 0.37 |
| **Oxygen** | $^{16}$O | 99.76 |
| | $^{17}$O | 0.04 |
| | $^{18}$O | 0.20 |

The uneven distribution occurs due to fractionation of the isotopes based on the mass difference. There are two types of fractionation effects that can occur, non-equilibrium and equilibrium. The non-equilibrium effect, also known as the kinetic effect, can happen when a reaction that is fast, incomplete or unidirectional occurs. Processes that can lead to kinetic effects include

diffusion, evaporation, biological reactions, as well as reactions that did not reach completion (Criss, 1999; Sharp, 2007; Wakefield, 2013).

Gaseous substances are said to all have the same average kinetic energy at a given temperature. Kinetic energy is calculated by Equation 2-1 below, where m is mass, and v is velocity. Two isotopologues of the same molecule have different masses and equal kinetic energy therefore different translational velocities. The heavier isotope has a lower velocity. This velocity difference can lead to kinetic fractionation effect for diffusion and evaporation processes. The lighter isotope with the faster velocity diffuses at a higher rate. In the case of evaporation, the lighter isotope has a higher chance of breaking through the surface layer as the energy binding it is lower, i.e. the molecule containing the lighter isotope will escape from the surface at a higher rate. In both cases, the residue ends up becoming enriched with the heavier isotope over time. Another process that can cause a kinetic fractionation effect is the difference in dissociation energies of molecules. For an incomplete reaction during the reaction rate step, if a bond breakage is needed in the rate-determining-step, the lighter isotope will be preferred (react at a higher rate) as less energy is required in order to break the bond. Figure 2-1 shows the lowest vibrational energies of hydrogen isotopologues H-H, H-D and D-D it is seen that H-H has the lowest vibrational energy which gives it the highest zero-point energy state and therefore less energy is needed for it to reach the state of dissociation compared to the other molecules, which leads to the residue of the reaction becoming enriched with the heavier isotope (Criss, 1999; Sharp, 2007; Wakefield, 2013).

$$K_e = \frac{1}{2}mv^2 \qquad \textit{Equation 2-1}$$

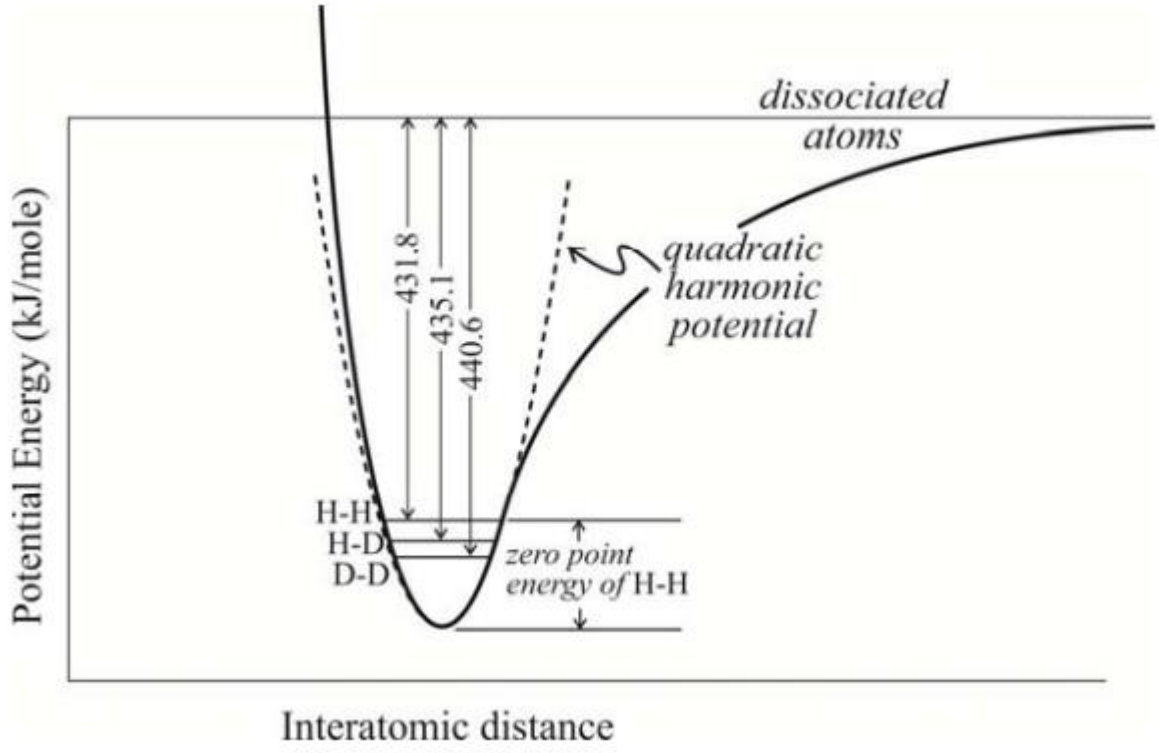


*Figure 2-1. Potential-energy curve for diatomic hydrogen. Showing that having at least one heavier isotope D decreases the zero-point energy and increases the amount of activation energy required to break the bond. Figure from Sharp (2007)*

Equilibrium fractionation effect occurs when a light isotope is substituted for a heavier isotope between phases containing the same element in a system. These fractionations tend to be smaller than that which is seen with the kinetic effect. The equilibrium fractionation effect has the most significant impact on hydrogen due to its high atomic mass ratio but also seen for carbon, oxygen and nitrogen. Modern mass spectrometers can detect such effects even in the heavy metals. (Criss, 1999; Sharp, 2007; Wakefield, 2013)

The atoms involved can be associated with chemical reactions (i.e. breaking and making bonds) or physical processes such as phase change. The vibrational energy related to these bonds is greatly affected by the slight change in mass difference associated with isotopes. An example of equilibrium fractionation effect is oxygen in water. Water containing $^{18}O$ is preferentially found in the liquid phase over the water-vapour phase when the system is at equilibrium (Criss, 1999; Sharp, 2007; Wakefield, 2013).

The isotope ratios of focus in the current study are $^{2}H/^{1}H$, $^{13}C/^{12}C$ and $^{15}N/^{14}N$. These are all classed as light stable isotopes. They all have a relatively low atomic number which gives more impact on the fractionation effects, as the proportion difference of the mass is greater. They have also been studied previously on plants for their use in discriminating based on geographical origin.

### *2.1.1 Hydrogen*

The hydrogen values in plants reflect the values of their source water, soil and/or irrigation. No fractionation occurs during the initial uptake of water via the roots. Thus, the xylem water in the roots and stem have the same $\delta^{2}H$ as the water in the soil. When the water moves from the stem into other plant tissues such as leaves and fruit kinetic fractionation can occur, as water evaporates out of the tissue leaving the tissue enriched with the heavier isotope (Ehleringer & Dawson, 1992). Given that soil is a good indicator of what the $\delta^{2}H$ values in plants are, it is important to understand how variation occurs in soil.

Soil water is a type of meteoric water which originates as atmospheric precipitation such as rain and snow. There is high variation across the globe for meteoric waters, which is due to the processes undergone by the atmospheric vapour including, evaporation, condensation and precipitation. Figure 2-2 shows the $\delta^{18}O$ and $\delta^{2}H$ of precipitated water along the global meteoric water line (Alexandre, 2020). This graph shows the isotope ratios are related to the regional temperature with an increase in $\delta^{18}O$ and $\delta^{2}H$ as the temperature increases.

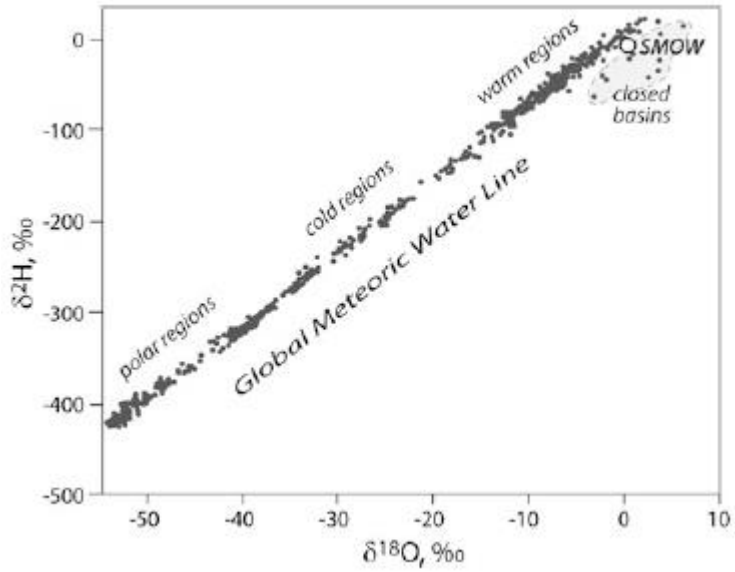


*Figure 2-2. Isotope values of precipitated oxygen and hydrogen from different regions. Producing the Global Meteoric water line. Figure from Alexandre (2020)*

Most atmospheric vapour originates as ocean water which has evaporated. As noted earlier, the lighter isotopes have a higher translational velocity which allows it to evaporate through the surface more easily. Therefore, the vapour starts more negative than the ocean source. Additional water vapour is added to the atmospheric vapour as it travels across the land from rivers, lakes and evapotranspiration in plants. Atmospheric vapour starts to cool down as it moves away from its source, which begins the process of condensation. Condensation of atmospheric water vapour follows the Rayleigh model shown in Figure 2-3 for $\delta^{18}O$, $\delta^{2}H$ of water vapour is correlated to $\delta^{18}O$ as seen in Figure 2-2 which makes this a good indication of the fractionation that occurs for hydrogen (Alexandre, 2020). Rayleigh fractionation model involves removal of a portion entirely from a reservoir, which changes the composition of the reservoir. In the new reservoir, this is repeated, and another portion is completely removed from the new reservoir. In the reverse of evaporation, the heavier isotope is preferred for condensation. As the heavier isotopes condense and precipitate out of the vapour the liquid water produced is removed from the system. This loss of water with an average $\delta^{18}O$ more positive than the vapour means the residual atmospheric vapour becomes increasing more negative. As the cycle progresses, the $\delta^{18}O$ of the vapour and the precipitation derived from it become increasingly more negative (Alexandre, 2020; Criss, 1999; Sharp, 2007).

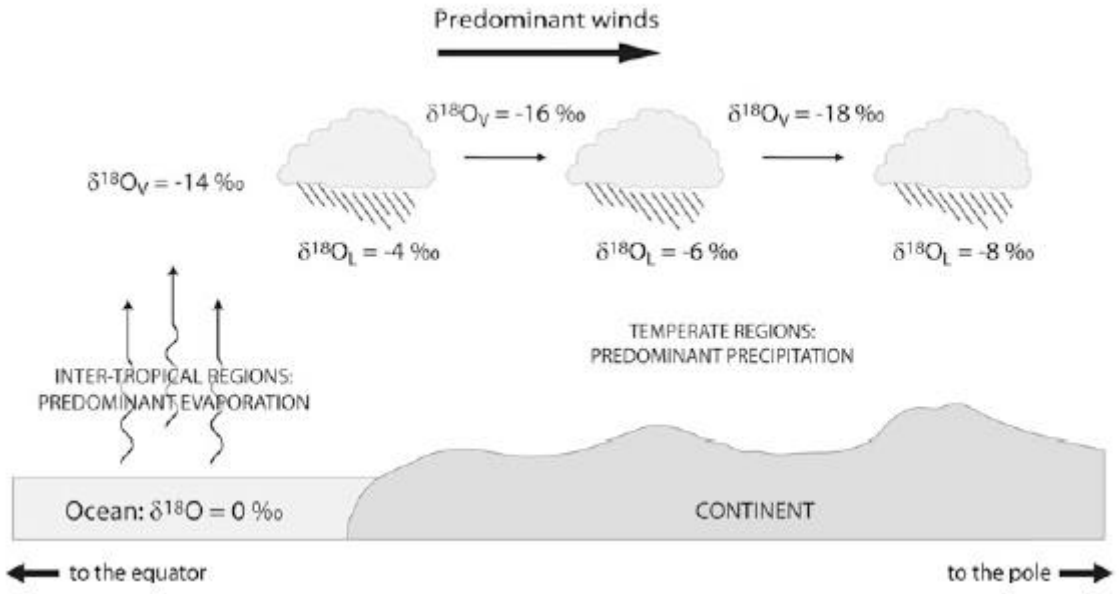


*Figure 2-3. Oxygen isotope values of the water evaporation and precipitation cycle. Figure from Alexandre (2020)*

Many factors are responsible for the variability in the precipitation that relates to the meteoric water. These factors often co-exist with one of the main factors being temperature. Temperature is the most important as it decreases the amount of water vapour that can be held in the air also decreases. As a greater proportion of the water condenses out, the atmospheric vapour will become isotopically lighter. This process leads to the observation of extreme isotopic depletion at the polar regions. Another factor that affects the $\delta^2H$ in atmospheric water is the continentality effect. As the atmospheric vapour moves further away from the source (predominantly tropical ocean surface) and goes farther in-land from the ocean, it becomes more negative. Temperature plays a role in this effect because as the vapour moves across the land, it will often move into colder climates. This effect also coincides with the latitude effect as the latitude increases the $\delta^2H$ values decrease. Again, this is predominantly due to the decrease in temperature and therefore, an increase in precipitation as latitude increases. Altitude also has a similar effect on the isotope ratio of atmospheric vapour as it becomes more negative as the altitude increases. Again, this is due to colder temperatures as altitude increases. The seasonal effect is also strongly linked to the temperature of the air as the rate precipitation is higher in colder months; this can also lead to more significant fractionation linked to the continentality effect as well (Criss, 1999; Sharp, 2007).

Lastly is the amount effect, is related to the amount of rain that occurs in tropical regions in different seasons. In months when there is a high amount of rain, the isotope ratio becomes

more negative compared to months with less rain. This is related to the fact that after it rains, the atmospheric vapor becomes more negative; therefore, as it continues to rain the precipitation will also become more negative over time. All these factors are related to the climate and geographical position of a region, which allows $\delta^2H$ of meteoric water to be a good indicator of origin (Criss, 1999; Sharp, 2007).

### *2.1.2 Carbon*

The carbon isotope ratio in plants is primarily driven by the photosynthesis process used by the plants. There are three different processes they can use $C_3$, $C_4$ and CAM (crassulacean acid metabolism). The photosynthesis process used affects the isotope values, as this is how the plant incorporates the atmospheric $CO_2$. In $C_3$ plants a carboxylation of ribulose bisphosphate occurs. Where $C_4$ the $CO_2$ undergoes carboxylation with Phosphoenolpyruvate, this it is then transferred by diffusion into a Bundle sheath cell where decarboxylation occurs. The $CO_2$ then undergoes carboxylation with ribulose bisphosphate. $C_4$ plants have been shown to have a less negative value of $\delta^{13}C$ with a mean value at around -13.5 ± 1.5 compared to $C_3$ plants which have a mean $\delta^{13}C$ value of around -28.1 ± 2.5 (Troughton as cited in O'Leary, 1981). This difference allows isotope ratios to be used as a way to identify if a plant uses a $C_3$ or $C_4$ process. Other plants use CAM to fixate their $CO_2$ these plants are found to have a broader range of $\delta^{13}C$ values. It has been found that CAM plants that live in a predominantly dark environment have less negative values than those in a light environment. The values for the CAM plants have been found to fall somewhere between the $C_3$ and $C_4$ plants (O'Leary, 1981). On a smaller scale, the $\delta^{13}C$ values in plants can be influenced by environmental factors such as light exposure, water stress and relative humidity as well as genetic variations. In plants, the $CO_2$ intake is regulated by the stomatal conductance and the photosynthetic demand of the plant. If the plant is under water stress, i.e. is unable to obtain sufficient water, this can cause the closure of the stomata to reduce the risk of losing water. The closure hinders the exchange of $CO_2$ from the atmosphere (J. F. Carter & Chesson, 2017). For $C_3$ plants when the stomata close it causes the plant cell to become a closed system. The exchange of the $CO_2$ with the atmosphere is stopped, but the photosynthetic process continues inside the cell. The photosynthesis leads to isotopic fractionation within the cell that can be described by a Rayleigh model. Carbon dioxide containing the lighter isotope reacts faster due to the lower activation energy required to break the bond leading to $^{13}C$ enriched reserves. As the process continues, the heavier isotope will be

utilised. Therefore the photosynthate product also becomes enriched (Marshall et al., 2008) leading to the $\delta^{13}C$ to increase in these plants.

The amount of light a plant is exposed to, which is essential for the photosynthesis process, can also cause change in the $\delta^{13}C$ of plant tissue. Plants grown in the shade have been seen to have a negative value compared to plants grown in sunlight (J. F. Carter & Chesson, 2017). With these environmental factors influencing the $\delta^{13}C$, it allows for the possibility of carbon values to be used to distinguish plants based on their origin environment.

### *2.1.3 Nitrogen*

Nitrogen isotope cycling in plants is more complex than that of carbon; hence nitrogen isotope values can be difficult to interpret. However, nitrogen isotope values in a plant are a good approximation to the nitrogen source. There are multiple processes in which a plant can uptake nitrogen which can also use different sources of nitrogen. Nitrogen from the soil can be up-taken via the plant's roots or a mycorrhizal fungal association. These two processes lead to the plant up taking nitrogen from the soil in the form of ammonium or nitrate. Both of these processes lead to the $\delta^{15}N$ in plants being dependant to the $\delta^{15}N$ of the soil. However, other processes for the uptake of nitrogen are not related to the $\delta^{15}N$ in soil. Plants that uptake biological nitrogen symbiotically through bacteria have $\delta^{15}N$ values which are found to be independent of that of the soil. (J. F. Carter & Chesson, 2017; Sharp, 2007).

Soil generally has a positive $\delta^{15}N$ value compared to the reference source of air. Nitrogen goes through many processes which can cause fractionations in soil. One such process is denitrification which forms $N_2$ gas through a two-step process from $NO_3^-$ and $NO_2^-$. This process is carried out by anaerobic bacteria, aerobic bacteria as well as some fungi. Denitrification can increase the $\delta^{15}N$ in soil as $^{14}N$ reacts faster than $^{15}N$, and the $N_2$ gas produced is lost to the atmosphere. The process of denitrification has the greatest effect on soils that are poorly drained and poorly oxidised (Sharp, 2007). Nitrification is another process that can cause some fractionation; it is the process of $NH_4^+$ being converted to $NO_3^-$ with $NO_2^-$ as an intermediate product. The first reaction is slow compared to the second part making it the rate-determining step and the step where the fractionation occurs. Nitrification leads to the $NO_3^-$ having a more negative $\delta^{15}N$ value than the $NH_4^+$ (Hoefs, 2015; Sharp, 2007). The $\delta^{15}N$ in soil

has also been linked to climate, with $\delta^{15}N$ values shown to decrease with increase in precipitation and decrease in temperature (Amundson et al., 2003). Another factor that can affect the $\delta^{15}N$ of soil is the amount of plant material in the environment, as plants and their leaves and fruits etc. decompose, they release their nitrogen back into the soil. Litter drop can increase the $\delta^{15}N$ of the soil as leaves and fruit are enriched in $^{15}N$ due to the fractionation that occurred at uptake. The more plants in an environment, the more of the depleted N is returned to the soil (Handley & Raven, 1992; Sharp, 2007).

Human activity also has an impact on the $\delta^{15}N$ of soil, with the most prominent impact coming from adding fertiliser to the soil. The $\delta^{15}N$ value of soil can reflect the type of fertiliser used as synthetic and organic fertilisers have different $\delta^{15}N$ values. Produced by the Haber-Bosch process for over 100 years, synthetic fertilisers are derived from atmospheric $N_2$ and $H_2$ to form ammonia ($NH_3$). Around 66 % of $NH_3$ is produced using this process (Kandemir et al., 2013; Milton et al., 2017). Due to atmospheric $N_2$ being used, the $\delta^{15}N$ for synthetic fertilisers is around 0 ‰. In comparison, a biologically based fertiliser has a more positive isotopic ratio due to the trophic level enrichment through the food chain (J. F. Carter & Chesson, 2017; Sharp, 2007).

Thus, while $\delta^{15}N$ can be used in authentication of plants; it is a product of local geology with a strong influence on farming practices, which can be used in authentication of origin based on the climate effects on the $\delta^{15}N$ in soil. However, it is more often used as a means to assess organic status based on the use of manure or synthetic fertiliser.

## 2.2 IRMS

Isotope ratio mass spectrometers (IRMS) are used to measure the isotope ratios of compounds in the form of simple gases CO, $CO_2$, $H_2$ and $N_2$ for O, C, H and N respectively. IRMS does not measure the quantitative amount; instead, it measures the ratio by simultaneously measuring the ion stream of the different isotope masses. There are three sections to an Isotope Ratio Mass Spectrometer; an ion source, a mass analyser and an ion collector (Benson et al., 2006; Meier-Augenstein, 2010).

In IRMS instrumentation the ion source describes the section where the sample in the form of a simple gas is transformed into an ion. The ion is formed due to the loss of an electron. Ionisation is achieved by the use of an electron beam impacting the molecule. Once formed, the ion is accelerated out of the ion source into the next section of the IRMS, the mass analyser. The mass analyser consists of a magnetic field that splits the ion stream up by the different mass-to-charge ratios m/z (Benson et al., 2006; Meier-Augenstein, 2010).

Each ion stream is sent to its specific collector, Faraday Cups. The Faraday cups (FC) collect the ion streams and simultaneously measure the isotope ratios. Each FC is placed in the exact position to be at the focal point of each ion stream. It is important to be able to set these up correctly, which is why simple gases are measured. The number of possible stable isotopes and the relative abundance for each are known for the simple gases. Therefore, it is easier to set up the collector to the correct masses. They also have a limited number of possible isotopomers which allows for only a few FC to be needed. For nitrogen and carbon, which can be measured simultaneously, three cups are needed. These are set up to collect m/z 28 and 29 for $N_2$ and m/z 44, 45 and 46 for $CO_2$. The use of FC in for the measurements of the isotope ratios of simple gases provides greater accuracy and precision in the results. The FC are not only set based on the m/z of the isotopologues but also their natural abundance. The FC are connected to individual amplifiers. The gain of these amplifiers reflects the abundance of the isotopologues in nature with the more dominant having a smaller relative gain to the rarer. As the ion stream runs through the FC, it is met with a resistor which converts the ion stream to a measurable voltage (J. Carter & Barwick, 2012; Criss, 1999; Meier-Augenstein, 2010).

When IRMS first came out, the gas introduction system used was the dual inlet system. This system required the sample preparation to be completed off-line, which required for there to be time and reasonable sample size for this to be completed. The off-line sample preparation consisted of breaking down complex compounds to the simple gas that is needed to measure the isotope ratios. One of the benefits of the duel inlet system was that it provides the means to measure both the sample and standard almost simultaneously, thus providing the best measurements of the ratio of the sample to reference composition. As technology advanced, a new introductory system was created, continuous flow, which is the most common method used today. The continuous flow system has a sample preparation step within it, such that the sample is turned into the simple gas inside the instrument itself, saving time and requiring a lot less of

the sample. The sample gases are then swept into the source of the mass spectrometer in a stream of carrier gas, usually helium (Meier-Augenstein, 2010).

Two types of sample preparation techniques were used with the continuous flow IRMS for the current study. These are Elemental Analyser (EA), and Gas Chromatography (GC) as to which technique is used is dependent on the sample type and what you want to measure. The EA is used to measure the isotope ratio values of the bulk sample. This means that if the sample were a complex mixture, the ratio measured would be the weighted average for the entire mixture. Where the GC method can be used to measure the isotope ratio of specific compounds within the mixture, this is because the GC is a separating technique and can separate the specific compounds within the sample before it is introduced to the IRMS. Both methods of continuous flow use a helium carrier gas, this helps move the sample through the instrument from introduction chamber right through to analysis in the Faraday Cups (Benson et al., 2006; Meier-Augenstein, 2010).

### *2.2.1 Elemental Analyser*

To measure the bulk δ values for carbon and nitrogen, the samples undergo combustion. For this process, a small aliquot (~0.8 mg) of the sample is weighed into a tin capsule. These capsules are closed and dropped into the IRMS using an automated sampler. The isotope ratios of N and C can be measured consecutively on the same sample run. The samples are combusted in a reactor with oxygen at temperatures between 900-1050 °C to form $CO_2$ and $N_2$, during this process $NO_x$ and $H_2O$ are also formed. The gases formed are then carried through to a reduction furnace. Here the $NO_x$ are reduced to $N_2$, and all excess oxygen is removed. The sample is then transferred through a drying tube to remove any water. Lastly, the samples are carried through a GC column in order to separate the $CO_2$ and $N_2$ before they are introduced into the IRMS for analysis (Benson et al., 2006; J. Carter & Barwick, 2012; Meier-Augenstein, 2010).

To measure the bulk isotope ratios for $\delta^2H$ and $\delta^{18}O$ which are run on using TC/EA (High Temperature Conversion Elemental Analyser), the sample undergoes high-temperature conversion. Like with nitrogen and carbon, hydrogen and oxygen can also be measured on the same run. The samples are weighed out into silver capsules and dropped into the reaction tube using an automated sampler. Here the sample undergoes a high temperature (1400 °C) pyrolytic

conversion to form the gases $H_2$ and CO. These gases are then separated through a GC column; this process also removes any interference for $N_2$ that may have formed, before they enter the IRMS (Benson et al., 2006; J. Carter & Barwick, 2012; Meier-Augenstein, 2010).

### *2.2.2 Gas Chromatography*

The GC continuous flow method is used to measure the isotope ratio of specific compounds extracted from the sample. A GC-IRMS has a capillary GC column which is used to separate the substance from others in the mixture. As long as the substances have different retention times through the GC, it is possible to analyse multiple substances within the mixture. Before a sample can enter the GC-IRMS, it needs to undergo some preparation to ensure that the target compounds are in a state volatile enough for GC separation, this usually includes extraction and derivatisation. It is then injected into the GC column, and the substances are separated by their different interactions with the stationary and mobile phases in the GC column. They are then carried into the reaction tube where they undergo either combustion for $\delta^{13}C$ and $\delta^{15}N$ or pyrolysis for $\delta^{2}H$ and $\delta^{18}O$. Unlike EA, $\delta^{13}C$ and $\delta^{15}N$ cannot be measured in the same run. For $\delta^{13}C$ analysis, once the sample undergoes combustion in an oxidation chamber $CO_2$, $NO_x$ and $H_2O$ can be formed; therefore, the $NO_x$ and $H_2O$ need to be removed. First, the $NO_2$ is reduced to $N_2$, and all excess oxygen is removed in a reduction chamber just as with the EA procedure. The $H_2O$ is removed in a nafion water trap. Then the sample is carried through to the IRMS. When $\delta^{15}N$ is being measured, the $CO_2$ must be removed as it can form CO in the source and cause an isobaric interference. The $CO_2$ removal is achieved by using a cryogenic trap (liquid $N_2$). Otherwise, the process is the same as that for $\delta^{13}C$ and EA (Benson et al., 2006; Meier-Augenstein, 2010; Muccio & Jackson, 2009).

For $\delta^{2}H$ measurement, the sample compounds are passed through an empty ceramic tube at 1450 °C where all the H is reduced to $H_2$. $\delta^{18}O$ analysis required the compounds to be passed through a ceramic tube containing platinum wire as a catalyst for conversion of the oxygen in the sample to CO. In both cases, the resulting gas, $H_2$ or CO, is then passed through a nafion $H_2O$ trap before entering the IRMS. Again, unlike with TC/EA, $\delta^{2}H$ and $\delta^{18}O$ cannot be measured in the same run (Benson et al., 2006; Muccio & Jackson, 2009).

### 2.2.3 Data Processing

The Faraday cups measure the ratio of the isotopes simultaneously. This means that the results are derived by comparing the ratios between the different isotopes rather than the absolute amount of each isotope in the sample. This procedure produces greater precision and smaller errors when comparing samples within the same analytical run. However, care must be taken in the calibration and quality control procedures to obtain results that are comparable between runs and between laboratories. The raw isotope ratios of the sample are compared to a standard for normalisation to the relevant international scale, the international standards often used for $\delta^{13}C$, $\delta^{15}N$, $\delta^{2}H$ and $\delta^{18}O$ are shown in Table 2-2 below. The ratios of the isotopes compared with the standard are recorded as delta values, δ with units of per mil, ‰, calculated by Equation 2-2. Calculating delta values can allow for better comparison within and between laboratories.

$$\delta = \frac{(R_{Sample} - R_{Standard})}{R_{Standard}} \quad \textit{Equation 2-2}$$

Where $R_{Sample}$ is the measured ratio of the sample and $R_{Standard}$ is the ratio of the standard. Both ratios are heavy to light isotope (Benson et al., 2006).

*Table 2-2. International Standards Used for C, O, N and H (Benson et al., 2006)*

| **International standard used** | **Isotope ratio** |
|---|---|
| PeeDee Belemnite (PDB) | $^{13}C/^{12}C = 0.0112372$ |
| PeeDee Belemnite (PDB) | $^{18}O/^{16}O = 0.0020671$ |
| Atmospheric nitrogen (AIR) | $^{15}N/^{14}N = 0.0036765$ |
| Vienna standard mean ocean water (V-SMOW) | $^{2}H/^{1}H = 0.00015576$ |
| Vienna standard mean ocean water (V-SMOW) | $^{18}O/^{16}O = 0.0020052$ |

In practice, each laboratory prepares in-house standards that are carefully calibrated to the international standards and run in each batch of samples for quality control purposes.

# Chapter 3 – Methods

Aim: Investigate if stable isotope measurements can be used to verify the origin of cocoa beans extracted from Chocolate.

To achieve the aim of this study and the objectives stated previously in section 1.6. Samples of cocoa beans from multiple known geographical origins (countries) and mastery of the analytical technique stable isotope analysis is required. This chapter outlines the samples and techniques used in this study.

## 3.1 Cocoa Beans – Samples

Eight bags of cocoa bean samples were obtained from Otago Chocolate Company (OCHO). Seven of the eight samples were of whole raw cocoa beans with the other being a bag of pre-roasted and crushed beans. Out of these samples, one sample originated in The Solomon Islands with the other seven from Papua New Guinea (PNG). Another eleven samples were provided by Whittaker's. Among the Whittaker's samples, there was one sample from Ghana and Samoa, four from Vietnam and five from Nicaragua. Individual raw cocoa beans from each sample bag were deshelled and manually ground up using a mortar and pestle. A portion of the pre-roasted sample was ground up using a mortar and pestle. Once ground, all samples were stored in individual sealed glass containers in a cool dark place. To be able to determine what the variation is within the bags of samples three individual beans were deshelled and ground for one bag from each of the different countries, expect for PNG where two different bags were explored.

## 3.2 Isotope Ratio Mass Spectrometry

The cocoa beans were analysed in a comparative analysis utilising the isotope ratios measured in both bulk cocoa beans and on individual fatty acids.

### *3.2.1 Bulk Isotope ratio analysis of cocoa beans*

The bulk $\delta^{13}C$, $\delta^{15}N$, $\delta^{18}O$ and $\delta^{2}H$ were measured for all cocoa bean samples. The C and N isotope ratios were analysed simultaneously using an EA-IRMS. About 0.8 mg of the ground up unshelled cocoa beans were weighed out in triplicate and placed into tin capsules, which were then folded to seal. The samples were then kept in a desiccator until analysis. The $\delta^{2}H$ as well as a second measurement of $\delta^{13}C$, were analysed using a TC/EA-IRMS. About 0.6 mg of each cocoa bean sample was weighed out in triplicate and placed into silver capsules. Capsules were then loosely closed to allow for equilibration with ambient water vapour to account for exchangeable H. These samples were left in the tray for at least ten days to equilibrate to the laboratory conditions. Equilibrating the samples to the laboratory conditions allows for the exchangeable H to match across all samples; therefore, the difference in results is due to the non-exchangeable H. They were then placed into a vacuum at 40 °C for four days to dry. Samples were then transferred to the autosampler tray and placed on a heat block under vacuum for one day before being run on the TC/EA-IRMS. An olive oil sample was measured every few samples to enable correction of instrumental drift.

### *3.2.2 Fatty acid analysis of cocoa beans – GC-FID and GC-IRMS*

To obtain the isotope ratios of the individual fatty acids in the cocoa beans, the lipids needed to be extracted and derivatised.

#### *Fatty Acid standards*

Two internal standards were used for the fatty acid analysis one with a smaller carbon chain and the other with a longer carbon chain than the fatty acids known to be in the samples, Nonanoic acid (C9) and Nonadecanoic acid (C19). These standards were chosen as neither are found naturally in the cocoa bean samples. The $\delta^{13}C$ and $\delta^{2}H$ values were determined for the two internal standards using the EA-IRMS and the TC/EA-IRMS, respectively. To measure the $\delta^{13}C$ around 0.4-0.5 mg of Nonadecanoic acid was weighed out in triplicate into tin capsules which were then closed and run on the EA-IRMS. About 0.6 mg of Nonanoic acid was weighed out using a syringe and placed into a smooth-walled tin capsule and sealed. The capsules samples were measured and placed into the EA-IRMS just before the individual run of the

sample to minimise possible losses through evaporation. The $\delta^2H$ values were measured by weighting out around 0.3 mg of Nonadecanoic acid into silver capsules which were then run. For the Nonanoic acid, the sample was injected directly into the TC/EA-IRMS via an autosampler. The $\delta^2H$ of the methanol was also measured this way. Unfortunately, it was not possible to get a consistent $\delta^{13}C$ value for the methanol due to it evaporating out of the capsule before it was able to be analysed.

An internal standard stock solution of nonanoic acid and nonadecanoic acid was made by weighing out about 0.1 g of each standard into a 10 mL volumetric flask. The flask was then filled to the mark with a 1:2 chloroform: methanol solution. The standard solution was kept in a sealed kimax tube in the fridge until used.

Standards for the four dominant fatty acids in cocoa beans were obtained as external standards. A solution of 10 mg in 10 mL was made up in a 2:1 chloroform: methanol solution for each standard. 1 mL from each solution was then combined to form a stock solution; this was made up in a 10 mL volumetric flask with the chloroform:methanol solvent. Solutions of the individual standard fatty acids, both external and internal were also made up. These samples were used to help determine the temperature programme for the GC along with the retention times for each fatty acid using a GC-FID. To do this, they underwent the same methylation process used on the samples as below.

### *Fatty acid extraction*

The extraction method used was based on the lipid extraction method by Bligh and Dyer (1959). 50 mg of the ground single unshelled cocoa beans were weighed out in triplicate and placed into separate screw-top kimax tubes along with 200 μL of the internal standard solution. 3.75 mL of a 1:2 Chloroform: methanol solution was added to the samples which were then vortexed for four minutes. An additional 1.25 mL of Chloroform was added then vortexed for two minutes. Finally, 1.25 mL of 8 % NaCl in distilled water was added to the sample and vortexed for an additional one minute. Samples were then centrifuged at 2000 g for five minutes. The bottom layer was removed and filtered using a pipette and glass wool into a second screw top kimax tube. The samples were then dried down in a stream of $N_2$ (g).

### *Fatty acid methylation*

The fatty acids were methylated to aid in the separation during analysis. 3 mL of a 6 % by volume concentrated $H_2SO_4$ in methanol solution was added to the kimax tubes with the dried extracted fatty acids and capped tightly. The samples were then placed in an oven at 80 $^{o}$C for two and a half hours, with the caps retightened after 15 minutes. Once out of the oven, they were left to cool for about ten minutes. Once cool to the touch, 2 mL of hexane was added, and the samples were vortexed for two minutes. 1 mL of distilled water was then added, and the samples were vortexed again for one minute, then left for one minute to allow the layers to separate. The hexane layer was then pipetted out and placed into 2 mL GC vials then capped. The samples were then run on the GC-IRMS using the TG225 column.

### *G.C. temperature programme – TG225 column*

The external standards were used to help determine the temperature programme for the GC. The methylated standards were run through a GC-FID at various temperature programs initially based off a temperature programme used by Mubiru et al. (2013). A Thermo Scientific TG-225MS column (Length 60 m, I.D. 0.32 mm, film 0.5 μm) was used. The final temperature programme was as follows. Oven temperature starting at 50 °C holding for 4 minutes, increasing to 225 °C at a rate of 12 °C$min^{-1}$ then holding maintaining at 225°C for a further 15 minutes.

### *GC-IRMS*

The methylated samples were analysed in groups of samples which underwent two different runs on the GC-IRMS to determine both the $\delta^2H$ and $\delta^{13}C$ values. A blank sample which contained only the two internal standards was run after every five samples. For carbon, a caffeine standard, made by weighing out about 0.05 mg of caffeine into a 2 mL GC vial with 2 mL of chloroform, was run at three different injection volumes to help account for any offset created by the reference gas value.

Before any samples can be run on the GC-IRMS, the mass spectrometer needs to undergo a stability run. Establishing the instrument is stable is essential before making measurements of the sample isotope ratios. The stability sequence involves the reference gas being measured ten times in a run, then comparing the values of each of these peaks. When the instrument is stable, the peaks will all be of similar isotope ratio value ($\sigma_{n-1} < 1$ for $\delta^2H$ and $< 0.1$ for $\delta^{13}C$). Therefore, ensure that the instrument is correctly picking up on the correct value for all samples in a run. The sequence is run as many times as it takes until the instrument is stable. Once the instrument is stable when measuring $\delta^2H$, the $H_3^+$ factor also needs to be measured. The measurement of the hydrogen isotope ratios is made in $H_2$ gas from the sample preparation. The simple gas $H_2$ is then ionised to form $H_2^+$ by electron impact in the source. However, when this ion is in the presence of a vast excess of $H_2$ and can create $H_3^+$, as shown in Equation 3-1 (Friedman as cited in Sessions et al., 2001). $H_3^+$ has mass three hence is indistinguishable from $DH^+$ in an IRMS, and the amount produced is dependent on the concentration of $H_2$ in the source. Therefore, a correction needs to be made this is very important as the $H_3^+$ can account for 5-30% of the signal (Sessions et al., 2001).

$$H_2^+ + H_2 \rightarrow H_3^+ + H^* \qquad \textit{Equation 3-1}$$

To calculate the $H_3^+$ factor, a sequence of ten reference gas peaks is run at varying concentrations of $H_2$ in the source. With the pressure starting to give a mass two signal at around 2000 mV and increased after every two peaks, the last two being at the highest pressure (highest concentration of $H_2$ in the source). The $H_3^+$ factor is then calculated as the ppm response per nA beam strength using a routine within the instrument programme. This factor is calculated daily and used for the entirety of the sample run.

### *Free Fatty Acid Column*

Extracted fatty acids of some of the cocoa bean samples were analysed on a second GC column (Agilent J&W CP-FFAP Length 25 m, I.D. 0.32 mm, film 0.3 μm) which allows for better separation of the fatty acids without having to methylate them. The methylation step adds a C and three H to the fatty acid. These additional C and H are then measured with the fatty acid on the GC-IRMS. Thus, additional calculations are required to determine the isotope ratio for the atoms from the original fatty acid. This correction increases the uncertainty around that final

result. Having a column that allows for the separation without the addition of the methyl group may, therefore, provide more accurate results. The samples were extracted as per the extraction method above. The dried fatty acid was then dissolved in 2 mL hexane and run on the GC-IRMS for $\delta^{13}C$.

### 3.3 Chocolate

The simplest form of chocolate is a mixture of cocoa beans and sugar. A couple of chocolate samples were made from the cocoa beans from Ghana. First, the cocoa beans were roasted in the shell at 120 °C for 40 mins, then let to cool to room temperature before the shells were removed. About 60 g of cocoa beans were weighed out and put into a blender with about 40 g of icing sugar. Icing sugar was used for this experiment instead of plain sugar as it added an extra component, tapioca starch, to the mixture making it more complicated. The mixture was blended until it became a liquid to simulate the grinding process of making chocolate. The mixture was then poured into a pestle and mortar and ground for about 20 minutes to simulate the conching stage. Then left to set covered in the fridge (*How to Make Chocolate From Scratch*, n.d.). When chocolate is created in a factory, the whole process usually takes days to complete, and it is also made on a much greater scale. The method used to create chocolate in this experiment simulates each process the beans go through on a small-time restricted scale. A second sample of chocolate was also made using this method with the addition of about 0.5 g of soy lecithin, which is an ingredient often used in chocolate.

The individual fatty acids of the chocolate samples were prepared using the same extraction and methylation method as used for the cocoa beans. These were then run on the GC-IRMS to determine the $\delta^{13}C$ and $\delta^{2}H$ values.

### 3.4 Work up of raw data

The data produced by the instrument requires some manipulation before it can be analysed. The raw results of a run can be affected by multiple factors. Factors include the stability of the instrument over the entire run. This can cause the instruments to drift in response throughout the run. This drift is the reason for measuring a standard at regular intervals through the run to

allow for correction. The standard used for this experiment was the internal standards, C9 and C19. These standards were included in each sample as well as the blank which was run, after every five samples. As these standards have a set ratio throughout the run, this can provide information on the instrument's performance.

For the $\delta^2H$ fatty acid standards, the methanol, C9 and C19 had been measured this allowed for the $\delta^2H$ value of the Fatty Acid Methyl Ester (FAME) to be calculated using Equation 3-2 below. Where $f_{FA}$ is the fraction of Hydrogens in the FAME that originate from the fatty acid and $f_{methanol}$ is the fraction from the Methanol.

$$\delta^2H_{FAME} = \delta^2H_{FA} \times f_{FA} + \delta^2H_{Methanol} \times f_{methanol} \quad \textit{Equation 3-2}$$

Solving this equation provides a known value of what the $\delta^2H$ for the two standards should be. That value can then be used to compare with the measured values to calculate for any drift. The samples were then adjusted to the drift of the C19. The reason C19 was used to calculate and adjust for the drift over C9 was that it was closer in size and retention time to the fatty acids that were being analysed. There was a noticeable issue with mass discrimination from the injector of the GC-IRMS. In which the later larger molecules were not leaving the injector chamber of the instrument entirely. This mass discrimination resulted in the C9 peak to be around twice the size of the C19 peak, where they should have been around the same height. It could also mean that some fractionation has occurred within the instrument itself, which will have an impact on the results. Mass discrimination is likely to affect the fatty acids of the samples similarly as they are of similar size to the C19. Therefore, making the corrections using the C19 will help to reduce the effect on the results created by the instrument.

Peak height is a crucial parameter to consider when it comes to measuring isotope ratios on the GC-IRMS. The ideal situation is that all working gas, standard and sample peaks are the same size as this generally leads to the highest precision of results. However, as the concentrations of sample peaks vary within and between samples, the ideal is unlikely to be achieved. An issue occurred during one of the runs for carbon, which was partially due to the greater size of the C9 peak, was that peaks were so high that in some of the runs they ended up topping out. Topping out meant that the isotope ratios provided for these runs were not meaningful. This

was another reason why the C19 was used to correct the samples as they were more reliable within this run. However, another issue faced in one of the hydrogen runs was that the instrument started to drift badly for the later samples. The drift is shown in Figure 3-1 below, where the samples from about position 60 in the run start to drop in value compared to the early samples. This drift makes it more challenging to adjust the sample values to ensure that what is analysed is an accurate result. The samples that were unable to provide satisfactory data after a drift correction were removed from any further analysis.

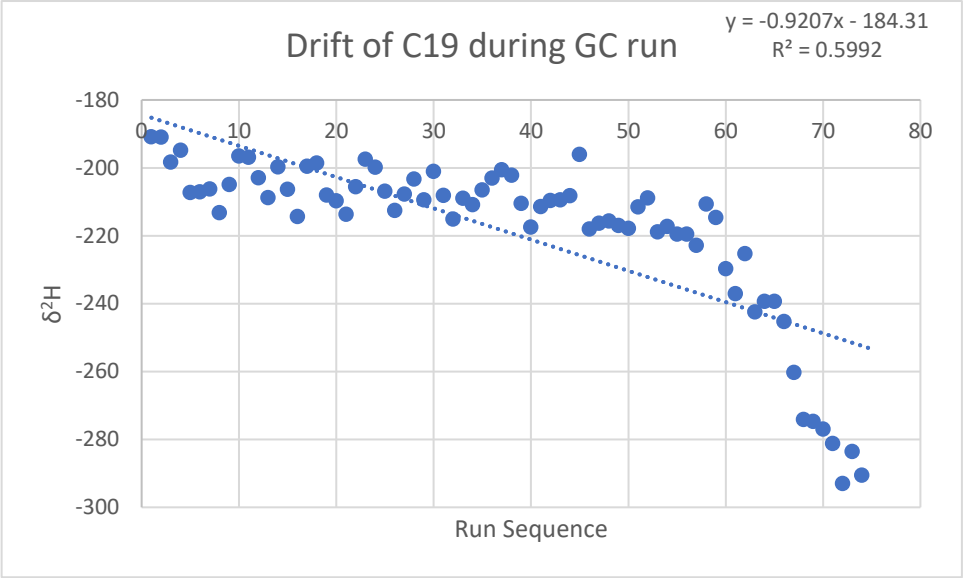


*Figure 3-1. Values of C19 during run of GC measuring $\delta^2H$ values. This figure shows the drift that occurred during the run with a major drop in values seen from about position 60.*

Once the data was adjusted for instrument errors, they were corrected for the addition of the methanol during the methylation process. The correction was achieved by rearranging Equation 3-2 above to solve of $\delta^2H_{FA}$ using the measured $\delta^2H_{FAME}$ for each sample and the known $\delta^2H_{methanol}$.

An issue observed for both the $\delta^2H$ and $\delta^{13}C$ runs was the poor separation of the peaks for C18 and C18-1. These two fatty acids both have very close retention times; this led to the peaks on the chromatogram not separating; instead, they showed two peaks with some overlap. This can be seen in the chromatogram from the Hydrogen run for a sample from Ghana (Figure 3-2). Due to the C18-2 fatty acid being of a much lower concentration than the other three fatty acids,

it did not always produce a peak that could be analysed. Therefore, this peak has been excluded from the data analyses.

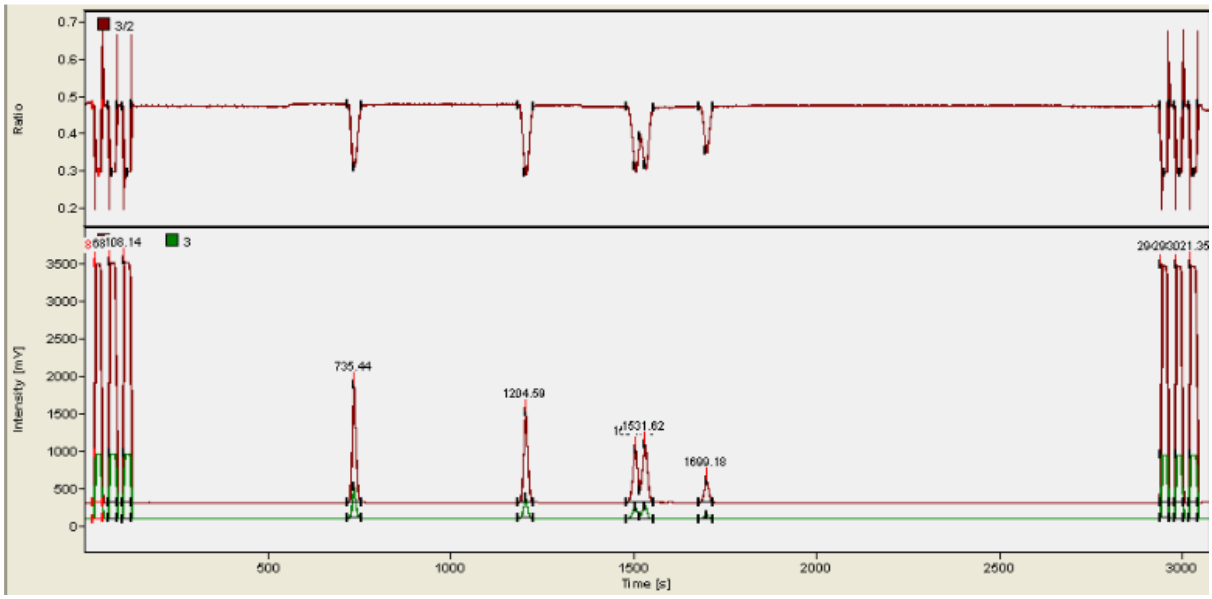


*Figure 3-2 Chromatogram for Hydrogen of a sample from Ghana, showing poor separation between peaks at 1504.8 and 1531.6*

The raw $\delta^{13}C$ values are calculated as the deviation from the value of the working (reference) gas. The reference gas is given a set ratio value in the instrument programme. This value is then used within the program to calculate the ratios of the samples. If this value is not the true and correct value of the reference gas, then an offset from the international scale will occur for all measured values. To normalise the raw data to the international scale, a caffeine sample with known $\delta^{13}C$ value was measured. The offset was calculated using Equation 3-3 below.

$$\Delta_{Caffeine} = \delta^{13}{}_{Caffeine-GC} - \delta^{13}{}_{Caffeine-IAEA} \qquad \textit{Equation 3-3}$$

Once the offset value was calculated, it was then applied to all the measure carbon results. After correcting for the offset the $\delta^{13}C$ for the methanol needed to be calculated. Due to the difficulty in obtaining the $\delta^{13}C$ of methanol via the EA, it was calculated using the measured C9 and $C9_{FAME}$ values. This was achieved by rearranging Equation 3-2 to give Equation 3-4 as below with the $\delta^{13}C_{FAME}$ equalling the average C9 value per run.

$$\delta^{13}C_{Methanol} = \frac{\delta^{13}C_{FAME} - \delta^{13}C_{FA} \times f_{FA}}{f_{Methanol}} \quad \textit{Equation 3-4}$$

The methanol $\delta^{13}C$ value was calculated using the data collected from all runs and averaged. The C9 value was used to determine the methanol over the C19 as the methanol added makes a more substantial portion of the C9 derivative. Therefore, the relative uncertainty in the calculation is smaller, providing a more reliable value. The methanol value was then used to calculate the values of the $\delta^{13}C$ of the FA for the samples and the C19 standard, by rearranging Equation 3-4 to solve for $\delta^{13}C_{FA}$. Lastly, the samples were corrected for instrumental fractionation. The correction was achieved by calculating the difference between the measured C19 in each run with the known value of the C19.

## 3.5 Statistical Analysis

Once all the raw data was corrected, it then underwent statistical analysis. This study aims to determine if it is possible to use stable isotope data to verify the origin of cocoa in chocolate. Statistical techniques used in this study include box plots to explore the data, and principal component analysis (PCA) was applied as a multivariate technique.

### *3.5.1 Initial exploration of bulk data*

Box and whisker plots were created for the $\delta^{15}N$, $\delta^{2}H$ and $\delta^{13}C$ bulk cocoa bean data to explore the distribution of the data to see if they showed any sign of differences between origin at the individual level. The spread of the cocoa bean results were analysed at three different levels. First within a bag of samples, then between bags of samples from individual countries and finally between the countries. These three levels of graphs were created to explore any variability seen between countries was not influenced by the variability seen between individual beans and bags. The box and whisker plot showed that there were some differences between countries. The next level of exploration was to see if combining variables help further distinguish between countries. Scatter plots were formed to compare pairs of variables ($\delta^{2}H$ vs $\delta^{13}C$, $\delta^{2}H$ vs $\delta^{15}N$ and $\delta^{13}C$ vs $\delta^{15}N$). These scatter plots showed some groupings between origins.

### *3.5.2 Initial exploration of fatty acid data*

The fatty acid results were first analysed with box and whisker plots of all three fatty acids (C16, C18 and C18-1) for both $\delta^{2}H$ and $\delta^{13}C$. These showed some differences between countries, so further analysis was conducted in the form of PCA.

### *3.5.3 Multivariate analysis: PCA*

The next step in exploring the data was to apply multivariate analysis, in the form of PCA as an unsupervised method to look for structure and clustering that can be attributed to the origin. PCA was applied to bulk ($\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$) and fatty acid ($\delta^{2}H$ and $\delta^{13}C$) data individually as well as a combination of both the fatty acid and bulk results.

Before the PCA was applied to the data, the data were standardised. This process meant that all variables were given a mean of zero and variance of 1. Standardisation is a necessary process as it brings all the variables onto the same scale. If standardisation is not conducted, then variables on a smaller scale may have their variances seen as insignificant and therefore not included in the analysis when, in fact, they could be very significant.

PCA is a multivariate analysis that is used to explain the majority of the variance between samples in a smaller number of variables than the original data. There are the same number of principal components as there are original variables. However, the new variables calculated for PCA are all uncorrelated from each other. They are also ordered such that the first variable contains the greatest amount of the variance, and the last variable contains the least. Therefore, even though there is an equal number of variables to the original data, the later variables will explain very little of the variances that they are not needed. Only the first few variables are used in a PCA as these variables will account for the majority of the variance (Manly & Navarro Alberto, 2016).

Principal components (Z) are calculated using the Equation 3-5. Where X is the original variables, p number of original variables and a is the loadings applied to ensure maximum variance of Z under the constants in Equation 3-6 are followed. For each Z calculated after the

first also have to ensure it is uncorrelated to all previous Z, i.e. for $Z_3$ needs to be uncorrelated to both $Z_1$ and $Z_2$ (Manly & Navarro Alberto, 2016).

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p \qquad \textit{Equation 3-5}$$

$$a_{i1}^2 + a_{i2}^2 + \cdots + a_{ip}^2 = 1 \quad \textit{Equation 3-6}$$

PCR was conducted using the R statistical programme with the packages ChemometricsWithR and Vegan.

# Chapter 4 – Results: Bulk Isotope Ratios

Isotope analysis was undertaken on the whole unshelled cocoa bean. The purpose was to determine if it is possible to use the isotope ratios of Carbon, Hydrogen and Nitrogen to differentiate the cocoa beans based on their country of origin. To achieve this differentiation, it was essential to understand different levels of possible variability within and between the samples. The variability at different scales needs to be taken into account when making a final interpretation of the data. If a sample has a low variability within one bag but a high variability across the multiple bags, this would show a high variability occurs between different batches of cocoa beans from across that country. However, if there is a high variability within one bag, it could have a more significant impact on the variability between multiple bags from the same country. The lowest level of variance is from the precision of the instrument. The precision value is a limitation of how well the instrument can measure repeatably. When calculating the standard deviations of the measurements, it is essential to consider this limitation. If values of calculated variances end up being less than the limit of the instrument, then this calculation is irrelevant as the precision is known to be the smallest possible variance. After the instrument precision, three further levels of variances were analysed. First, the standard deviations of samples which were determined from within a single bag. Next, the spread between different sample bags originating from the same country was considered. Finally, the variability between samples from different countries. Once these variances were analysed, a Principal Component Analysis (PCA) was conducted on the isotope values to determine if it was possible to differentiate the cocoa beans based on country of origin.

Aim: Investigate if stable isotope ratios of whole cocoa bean reflect origin.

Objectives: Determine if stable isotope ratios do reflect the geographical origin of cocoa beans

1. Explore variability of isotope ratios within a sample bag
2. Explore variability of isotope ratios within a country
3. Explore variability of isotope ratios between country
4. Evaluate any clustering of origins based on statistical analysis.

### 4.1 Variability within a bag

The variability within a single bag for $\delta^{13}C$ and $\delta^{2}H$ was analysed for all countries. One bag was randomly picked for each country, and three or four beans per bag were analysed. Except for The Solomon Islands where only two beans were analysed. For $\delta^{15}N$ the variability within a single bag was analysed for Nicaragua, Papua New Guinea (PNG) and Vietnam. Figures 4-1a, 4-1b and 4-1c show the mean of the samples from the bag with the standard deviation as the error bars, the means and standard deviations are also given in Table 4-1.

*Table 4-1. Mean δ and Standard deviation of samples measured from within a single bag for $\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$*

| | | $\delta^{2}H$ | | $\delta^{13}C$ | | $\delta^{15}N$ | |
|---|---|---|---|---|---|---|---|
| | Number of samples | Mean | Standard deviation | Mean | Standard Deviation | Mean | Standard Deviation |
| Nicaragua | 3 | -107.61 | 7.88 | -27.23 | 0.23 | 5.11 | 0.76 |
| PNG-1 | 4 (3 for N) | -160.10 | 11.85 | -29.06 | 0.45* | 3.33 | 0.93 |
| PNG-2 | 4 | -150.83 | 4.37 | -28.63 | 0.45* | | |
| Vietnam | 3 | -119.92 | 8.66 | -28.02 | 2.33 | 4.10 | 0.59 |
| Samoa | 3 | -137.67 | 2.01* | -29.28 | 0.72 | | |
| Solomon Islands | 2 | -131.87 | 1.08 | -30.81 | 1.30 | | |

*Precision value used to calculate standard deviation as variance was less than the analytical precision value; therefore, population variance irrelevant.

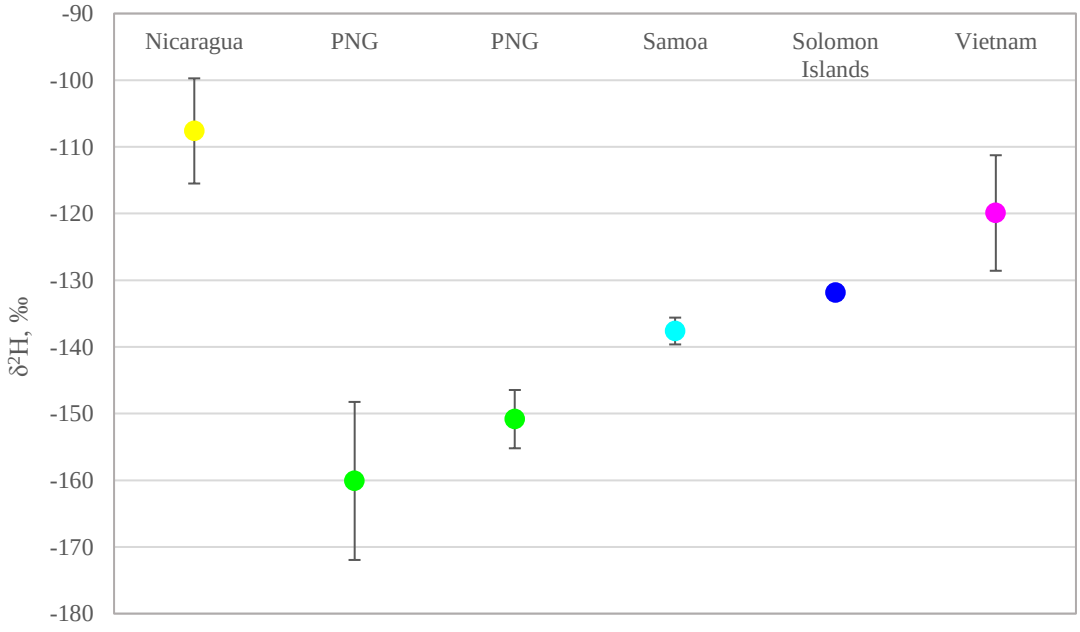


*Figure 4-1a. $\delta^{2}H$ averages of cocoa bean samples from within a sample bag, with standard deviation error bars.*

Figure 4-1a shows the mean and standard deviation of the $\delta^2H$ values for each country. From this graph and data given in Table 4-1, it can be seen that the mean values for each sample are different. The standard deviations also show a range of values. The variability within two different bags from PNG was analysed. One of these samples is seen to have the most significant standard deviation of 11.85 across all the countries. The second sample from PNG showed a much smaller standard deviation of 4.37. This shows that the variability within a bag can differ between countries as well as within the same country. The standard deviations shown are calculated from a small sample set; therefore, provides an indication of the variability within the bag and not an exact value. The smallest standard deviation was seen for the samples from The Solomon Islands, which had a standard deviation of 1.08. However, this low standard deviation may be coincidental due to only two cocoa beans being measured for this country. Samoa also showed a relatively small standard deviation of 2.01. The standard deviation for Nicaragua and Vietnam were larger at 7.88 and 8.66, respectively.

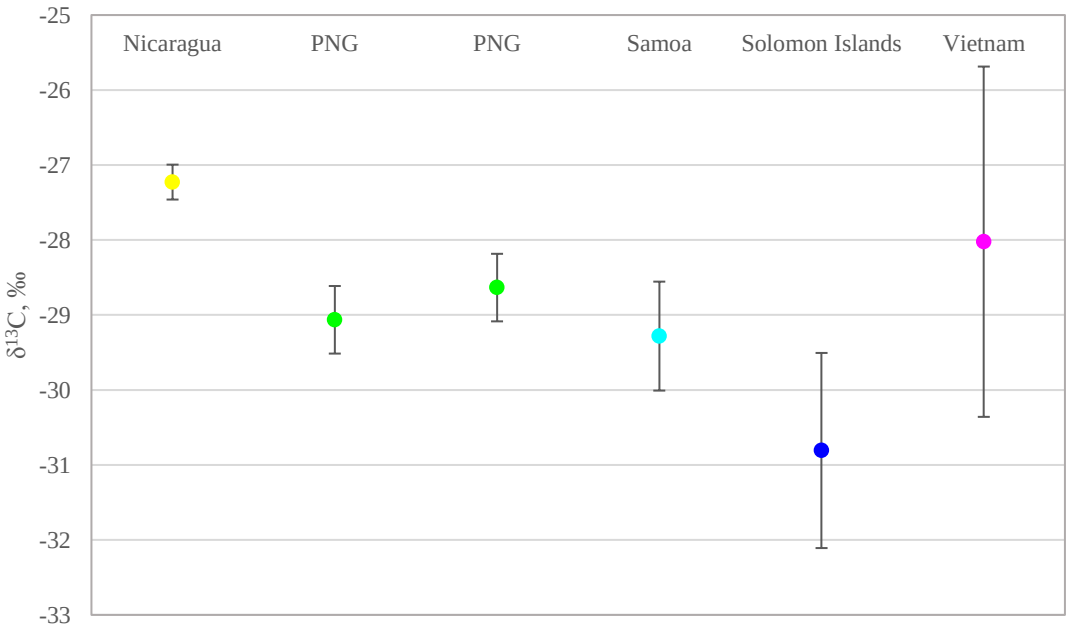


*Figure 4-1b. $\delta^{13}C$ averages of cocoa bean samples from within a sample bag, with standard deviation error bars.*

Figure 4-1b is the graphical representation of the mean and standard deviation within a single bag for the $\delta^{13}C$ values. This graph also indicates that there is some difference in the means between the different countries. The standard deviations (Table 4-1) also show a broad range of values. Notably is the standard deviation of Vietnam, 2.33. This is seen to be much larger than the standard deviation of all the other countries. This difference may be due to the way the

beans have been fermented. When looking at the physical appearances of the three beans analysed from Vietnam, they all appeared to be a different colour. The brown colour that is often associated with chocolate and cocoa beans is developed through the fermentation process. During the fermentation process, many different chemical reactions occur within the bean in which $CO_2$ is released as a by-product (Ziegleder, 2009). These reactions can have an effect on the $\delta^{13}C$ values as it involves the loss of some carbon. That partitioning of carbon may cause fractionation in the carbon isotopes. Therefore, if one bean had not been fermented as well as another, there may be a difference imparted in the carbon isotope values. The standard deviations of the other countries are much smaller. Samoa and The Solomon Islands are seen to have the next largest standard deviation of 0.72 and 1.30 respectively. The other countries all have very low standard deviations of 0.23 and 0.45, for Nicaragua and PNG respectively.

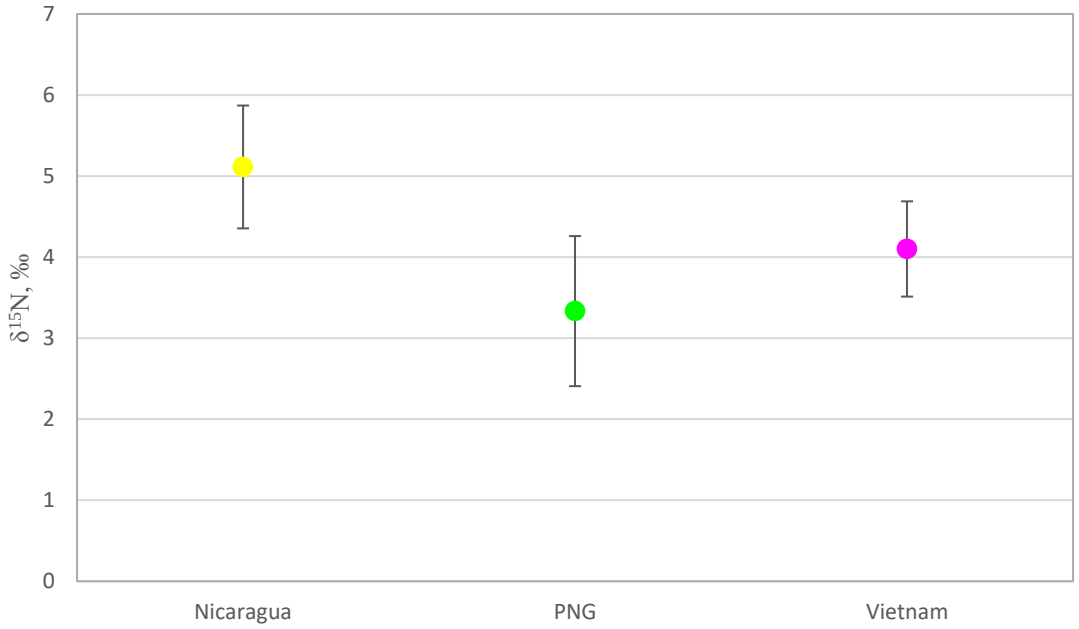


*Figure 4-1c. $\delta^{15}N$ averages of cocoa bean samples from within a sample bag, with standard deviation error bars.*

The standard deviations for $\delta^{15}N$, as seen in Table 4-1 and Figure 4-1c show that Vietnam had the lowest standard deviation of 0.58, and PNG had the highest standard deviation of 0.96. Nicaragua has a standard deviation between these at 0.75. Nitrogen isotope ratios were analysed on a different instrument to both the Carbon and Hydrogen values. The Nitrogen results were also found to have the lowest precision of the three. This low precision means that the measured values more closely represent the true value of the sample.

From these three graphs, it can be seen that the $\delta^2H$ values had the greatest difference in means and standard deviation values. Hydrogen also had the worst instrument precision. With the greater range in values for Hydrogen for these few samples, it can be seen that there is a difference in the bag means, which means that there is the possibility to use the Hydrogen isotopes to separate the cocoa beans by origin. Carbon and Nitrogen also show this possibility. However, there was greater overlap in these values. The biggest concern for the carbon values is the sample from Vietnam. The standard deviation indicates possible issues in separating these samples as they are seen to overlap with most of the other samples. A limitation of this analyse was the small sample size as only three or four beans were analysed per bag. This could mean that the standard deviation shown may not be a true representative of the actual standard deviation for each bag. Analysing a larger sample set would give better results. However, it would also mean more time and cost, which is not always possible or practical.

### 4.2 Variability within a country

For Nicaragua, PNG and Vietnam, multiple bags of cocoa beans were provided. Therefore, it was important to understand what variability there was between the different bags from these countries. Multiple beans were analysed per bag. Box and whisker plots were created to analyse the variability within the country, Figures 4-2a, 4-2b and 4-2c. It can be seen in all three graphs that the mean value (x) for the samples is very close or equal to the median value.

Figure 4-2a shows the $\delta^2H$ variability of the three countries. It can be seen that PNG has a relatively small range within the upper and lower quartiles but has larger difference between the maximum and minimum values and therefore, greatest spread over all the samples. Where the samples from Nicaragua have a greater range between the upper and lower quartiles but shows very little to no difference from the upper and lower quartile values to the maximum and minimum values respectively. Indicating the samples at the ends have very similar values. Vietnam samples show that similar range between the upper and lower quartiles as Nicaragua. Additionally, there was nearly no difference to the lower quartile and minimum value. However, there is a difference seen between the upper quartile and the maximum value.

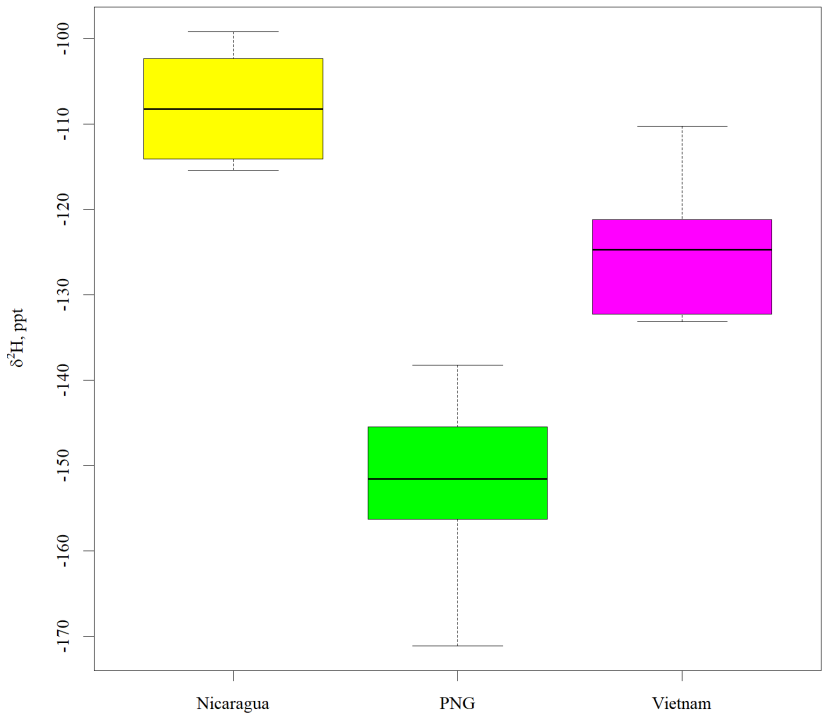


*Figure 4-2a. Box and Whisker plot showing the range of $\delta^2H$ bulk cocoa beans within the countries, Nicaragua (7), PNG (15) and Vietnam (6).*

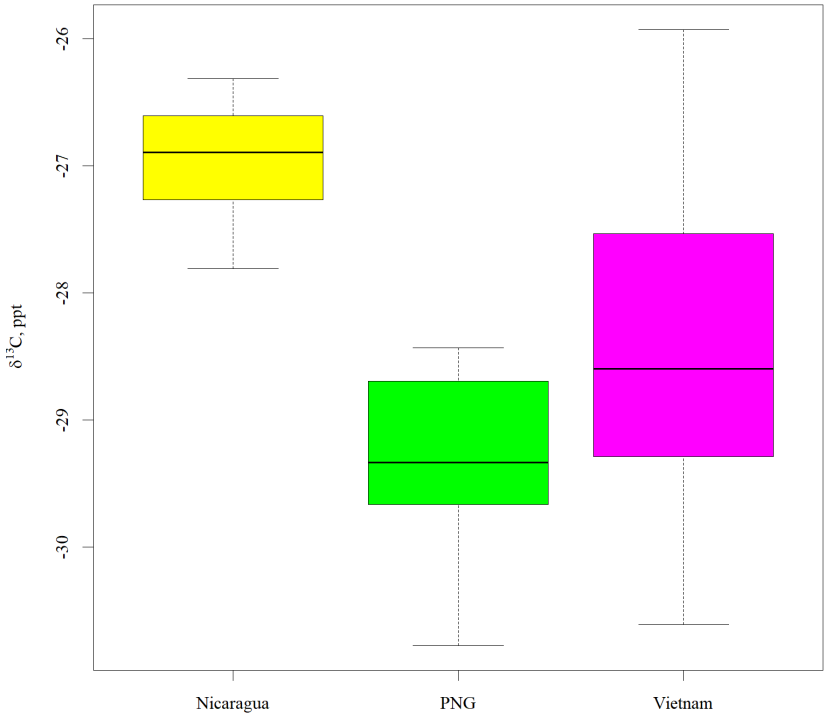


*Figure 4-2b. Box and Whisker plot showing the range of $\delta^{13}C$ bulk cocoa bean samples within the countries, Nicaragua (7), PNG (15) and Vietnam (6).*

Box and whisker plot of the $\delta^{13}C$ (Figure 4-2b) shows that Nicaragua exhibits the least difference between the different sample bags. It can also be seen there is more spread in the samples which have values higher than the mean compared with the samples with lower values. Where PNG has a small range between the median and the lower quartile relative to the upper quartile. However, there is a larger spread between the lower quartile and the minimum value compared with the upper quartile and maximum value. The spread between the samples from Vietnam is significantly larger than the spread for both PNG and Nicaragua. This was also seen when looking at the standard deviation from within a bag Figure 4-1b. Therefore, the variability observed here may be affected by the variability within a bag.

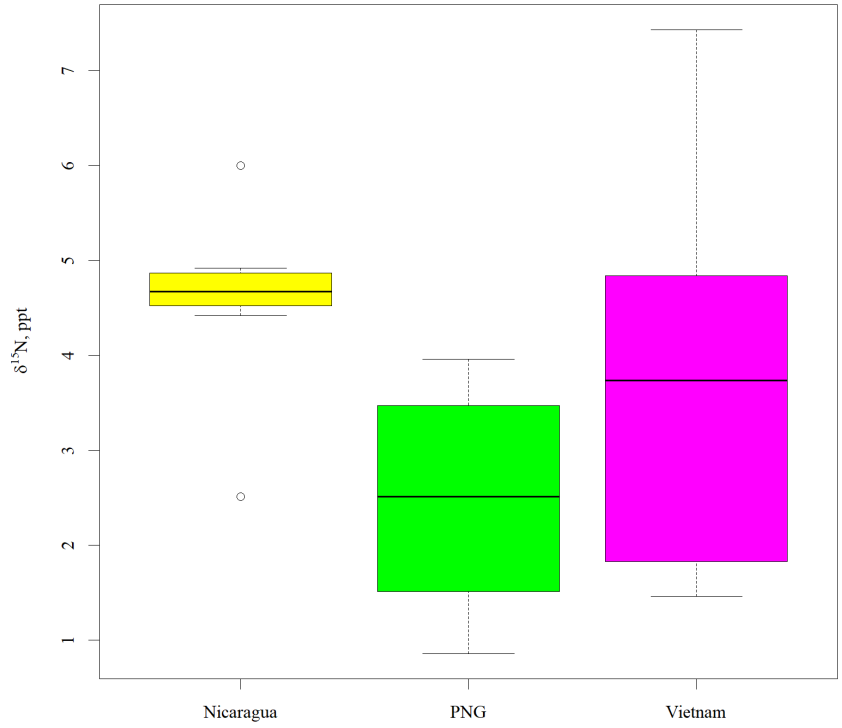


*Figure 4-2c. Box and Whisker plot showing the range of $\delta^{15}N$ in bulk cocoa beans within the countries Nicaragua (7), PNG (8) and Vietnam (6).*

The variability between the different bags for $\delta^{15}N$ Figure 4-2c also showed Nicaragua as having the smallest range of the three countries. With most of Nicaragua samples being very close together in value. Expect for two outlier samples at the maximum and minimum. Vietnam showed the largest range between the bags. With there being a larger range between the median value and the maximum value. Unlike with the carbon values, the variability within a single bag of Vietnam showed a much smaller range for the $\delta^{15}N$ (Figure 4-1c). This indicates that the variability between the bags is not going to be related to the within in a bag variability as with the carbon values. Therefore, this difference is influenced by geography and processes used on the beans. PNG also show a relatively large range for $\delta^{15}N$ compared to $\delta^{2}H$ and $\delta^{13}C$.

## 4.3 Variability between countries

After analysing the variability within a single bag and between multiple bags from the same country, an analysis to see if there was any variability between the different countries was conducted. Box and whisker plots were created for each isotope ratio to determine if there was any observable difference between the countries (Figures 4-3a, 4-3b and 4-3c). The mean δ values for each country can be seen in Table 4-2.

*Table 4-2. Mean δ value of H, C and N for each country*

| | $\delta^2H$ | | $\delta^{13}C$ | | $\delta^{15}N$ | |
|---|---|---|---|---|---|---|
| | Number of samples | Mean | Number of samples | Mean | Number of samples | Mean |
| Ghana | 1 | -116.23 | 1 | -28.39 | 1 | 6.15 |
| Nicaragua | 7 | -107.95 | 7 | -26.97 | 7 | 4.56 |
| PNG | 15 | -152.20 | 15 | -29.28 | 8 | 2.48 |
| Samoa | 3 | -137.62 | 3 | -29.28 | 1 | 3.27 |
| Solomon Islands | 2 | -131.87 | 2 | -30.81 | 1 | 5.66 |
| Vietnam | 6 | -124.39 | 6 | -28.43 | 6 | 3.84 |

For the $\delta^2H$ values (Figure 4-3a), there was some difference seen between the countries, with all countries showing different mean values. This difference is especially noticed when comparing Ghana and Nicaragua with PNG and Samoa. The $\delta^2H$ values range from around -100 to -170 per mil. The $\delta^{13}C$ values (Figure 4-3b) also show differences between the countries with a range of values from around -26 to -32 per mil. Most of the means for each country are different, except for PNG and Samoa, which both have a mean of -29.28. The greatest variability is seen between the samples from Nicaragua and The Solomon Islands. For the $\delta^{13}C$ values, the large spread within the samples from Vietnam means that this one country is showing a range of values which makes it unable to be separated from all the other countries by this parameter alone.

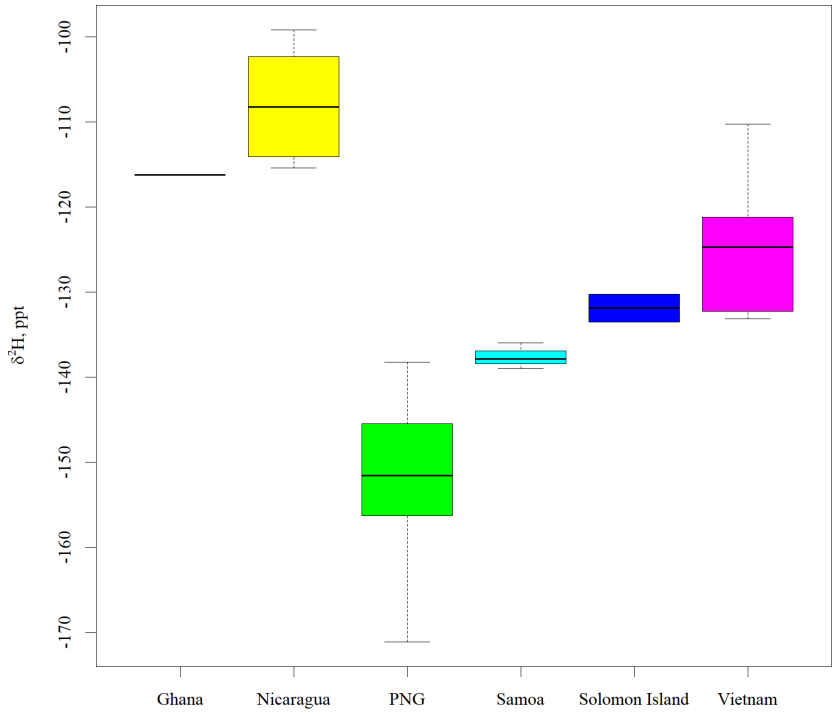


*Figure 4-3a. Box and Whisker plot comparing $\delta^2H$ values of bulk cocoa beans H across all countries. Ghana(1), Nicaragua(7), PNG(15), Samoa(3), Solomon Island(2), Vietnam(6)*

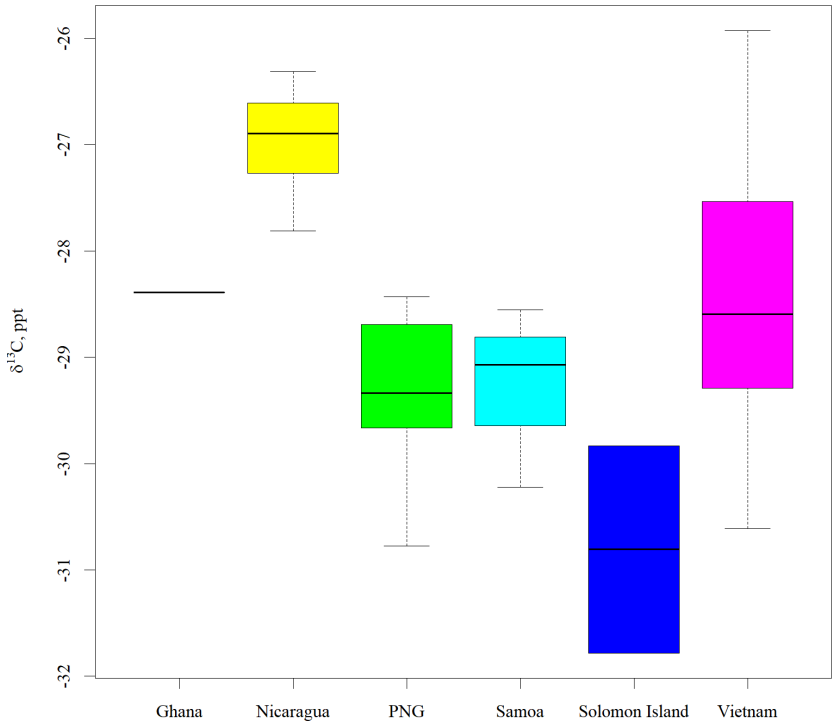


*Figure 4-3b. Box and Whisker plot comparing $\delta^{13}C$ values of bulk cocoa beans across all countries. Ghana(1), Nicaragua(7), PNG(15), Samoa(3), Solomon Island(2), Vietnam(6).*

The $\delta^{15}N$ values (Figure 4-3c) show a range of around 1.0 to 7.5 per mil. The country means are all different from each other. Vietnam has a range of values that spread over all the other countries. The biggest difference is seen when comparing the means of Ghana and The Solomon Islands with PNG and Samoa. These three graphs show that there is some difference in the isotope values based on the country of origin. However, the sample size that has been analysed is very small, with some countries only having one sample. Therefore, as a base, this analysis shows a difference between countries, but to specify what the true isotopic value is for all of these countries, more samples will need to be measured.

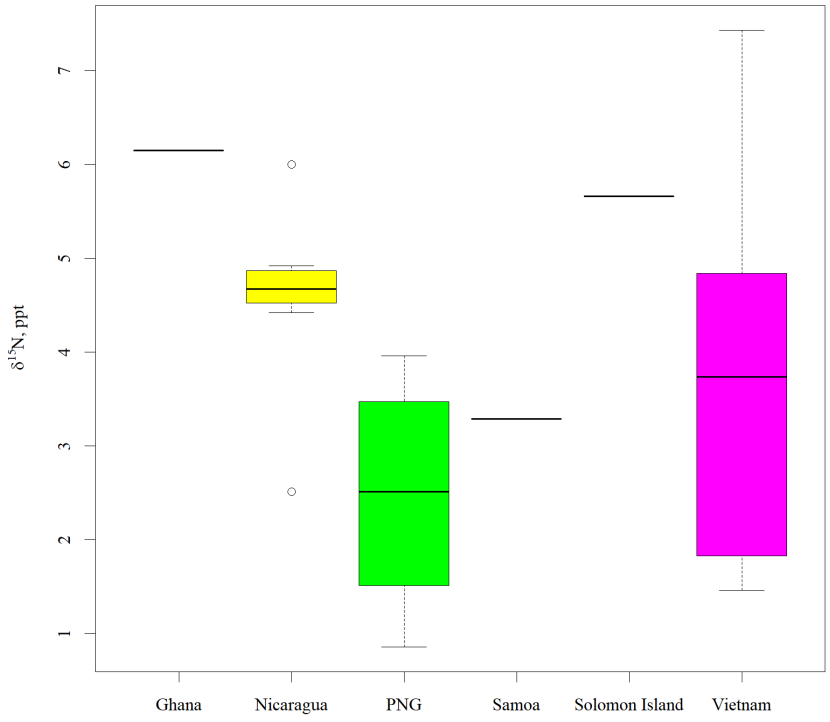


*Figure 4-3c. Box and Whisker plot comparing $\delta^{15}N$ values of bulk cocoa beans across all countries. Ghana(1), Nicaragua(2), PNG(8), Samoa(1), Solomon Island(1), Vietnam(6).*

While for $\delta^{13}C$, the high range is possibly due to the fermentation process of the samples. The range in the $\delta^{15}N$ may be due to the geography of the origins. It is known that some of the samples from Vietnam originate in the Central Highlands where others originate in the south, including within an area which has a maze of rivers flowing through. The variety of samples from Nicaragua are all maintained by the same supplier. So even if they are from a slightly different regions, they may undergo similar processes, which could explain why there wasn't a large range seen within this country.

## 4.4 Multivariate analysis of Bulk Isotope Ratios

The previous graphs showed that there are some differences seen between countries for each isotope. To determine if it was possible to separate the cocoa beans by country, multivariate analysis was applied firstly combining two isotope variables together to see if any patterns occurred and then analysing all variables in a PCA.

The isotope values were plotted against one another shown in Figures 4-4a, 4-4b and 4-4c. In the first plot Figure 4-4a, it compares the $\delta^2H$ values with $\delta^{13}C$. From this graph, it can be seen that two groups form one group containing samples from PNG, Samoa and The Solomon Islands, the other Nicaragua and Ghana. The samples from Vietnam are seen to be spread out across both groups. This is likely due to the large range in $\delta^{13}C$ seen within this country.

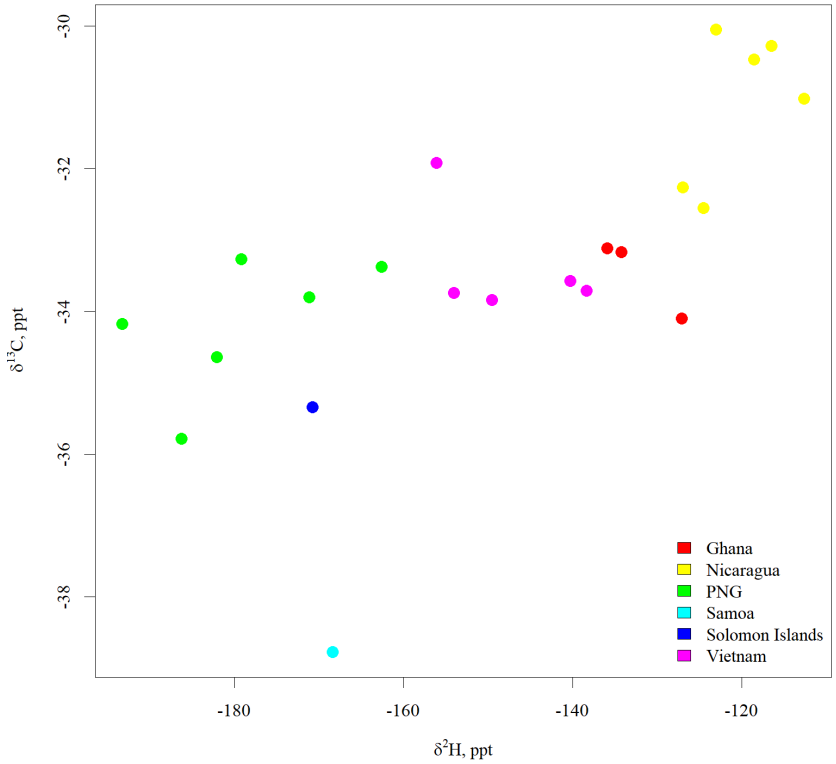


*Figure 4-4a. Plot of $\delta^{13}C$ vs $\delta^2H$ values of bulk cocoa beans, each country represented as different colour.*

The $\delta^{13}C$ and $\delta^{15}N$ values are compared in Figure 4-4b. These graphs also show a distinct separation between the samples from Nicaragua and Ghana with Samoa and PNG. The samples from The Solomon Islands can distinctly be seen as separated from the other countries. The samples from Vietnam are again showing a range of values that is spread across all the other countries. Lastly, the $\delta^2H$ and $\delta^{15}N$ values were compared, as seen in Figure 4-4c. This graph

shows that the samples from PNG are separated from the other groups. Ghana and Nicaragua are seen to have similar values again, grouping them close together. Ghana and Nicaragua together and PNG and Samoa together. The sample from The Solomon Islands is seen out on its own. Due to the large range in $\delta^{15}N$, samples from Vietnam can be seen to spread out across the graph. These three graphs show that there are some distinctive groupings among the samples. These groupings indicate that it can be possible to separate out the cocoa beans based on origin. In order to determine if it is possible to separate out all these countries from each other, a PCA was performed using all three Isotope values.

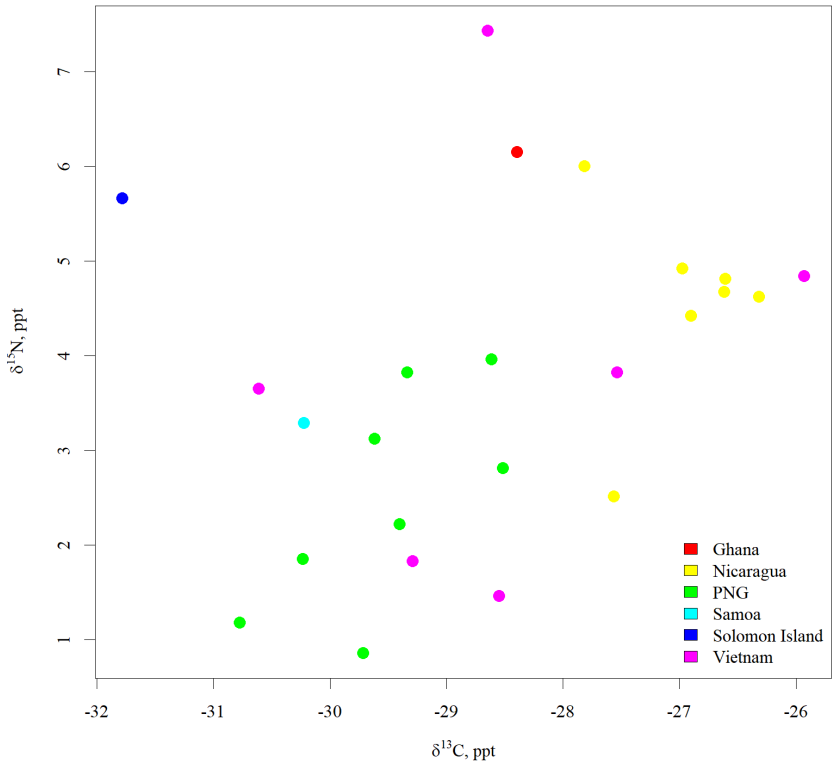


*Figure 4-4b. Plot of $\delta^{15}N$ vs $\delta^{13}C$ values for bulk cocoa beans, each country represented as different colour.*

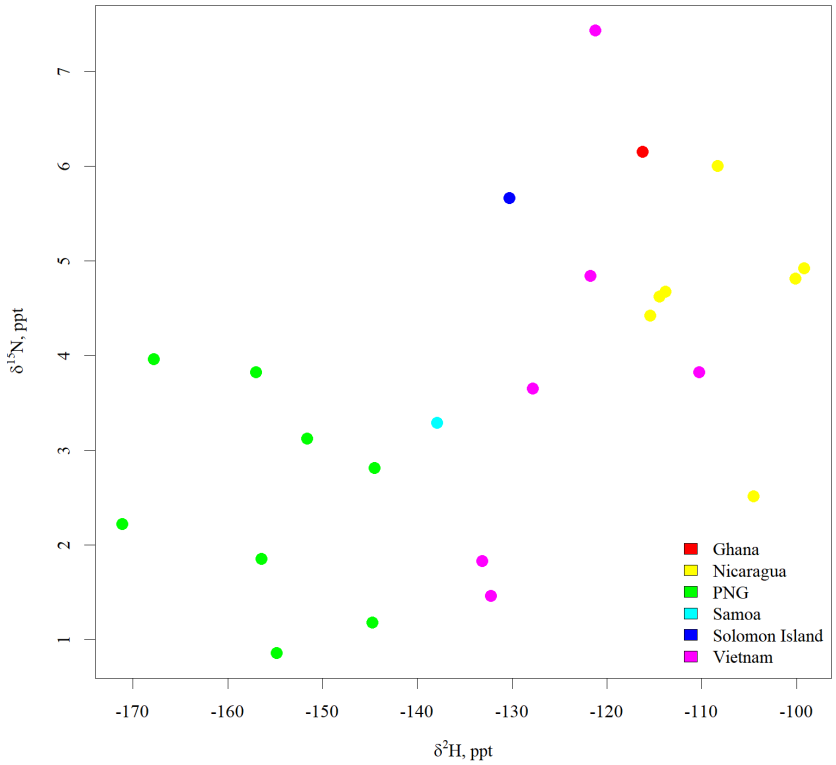


*Figure 4-4c. Plot of $\delta^{15}N$ vs $\delta^{2}H$ values for bulk cocoa beans, each country represented as different colour.*

Two PCA biplots were created using all three variables to see if it was possible to separate the countries from each other. The first PCA biplot (Figure 4-5a) compares the first two Principal components, PC1 and PC2. From this graph, two distinct groups can be seen. One group containing samples from PNG and Samoa. The other with samples from Nicaragua with the sample from Ghana close by. Both groups also contain samples from Vietnam, which as previously stated has been seen to have a large range between the samples. The sample from The Solomon Islands can be seen to be completely out on its own. The second PCA biplot which plots PC1 with PC3 (Figure 4-5b) still shows two distinct groups one of Nicaragua and Ghana and the other of PNG, Samoa and The Solomon Islands.

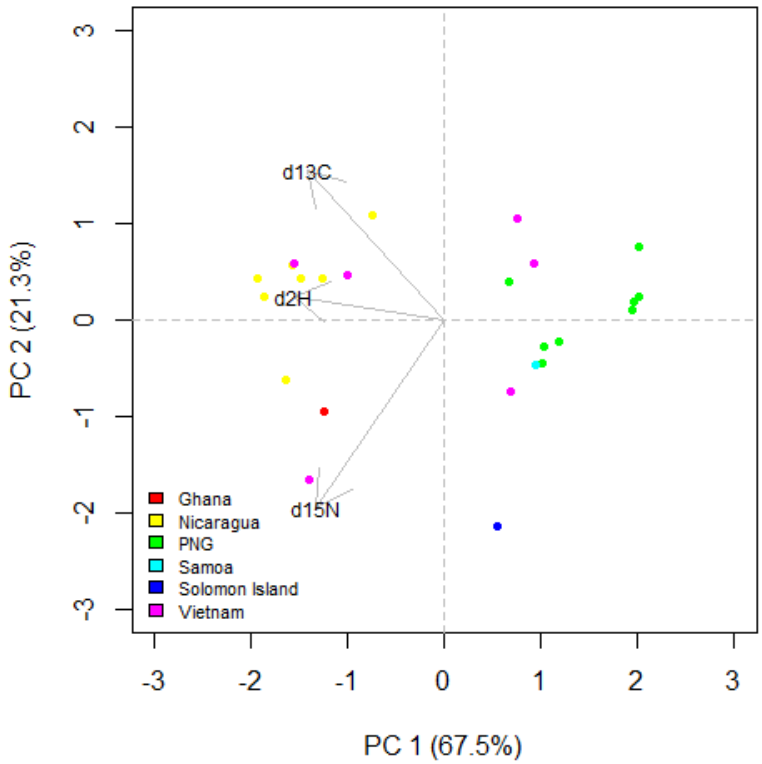


*Figure 4-5a. Biplot of PC1 vs PC2 for PCA on bulk $\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$ of cocoa beans from Ghana, Nicaragua, PNG, Samoa, Solomon Islands and Vietnam.*

Comparing graphs shows that PC3 is needed to improve the separation between countries. this is especially seen for the samples from Samoa. These samples almost indistinguishable in the PC1 vs PC2 biplot (Figure 4-5a) to samples from PNG, but are separated in the PC1 vs PC3 biplot (Figure 4-5b). From this analysis, it is possible to use isotope ratios to separate the cocoa beans based on their country of origin. The only exception being for Vietnam whose large range within and between samples has meant that as a country it does not group together; rather, it is shown to spread out over groups of other countries. This analysis also has its limitations as the sample size is very small, especially for the countries Samoa, The Solomon Islands and Ghana in which only one sample has been analysed. Increasing the sample size will help to know if the samples analysed here are the true value for the country. When you are only analysing one sample from a country, all you really can say is that this sample is different from the other countries. Also collecting multiple samples from multiple regions of a country would help know the true range within the country and if the spread seen here is due to the regions in which the samples have originated.

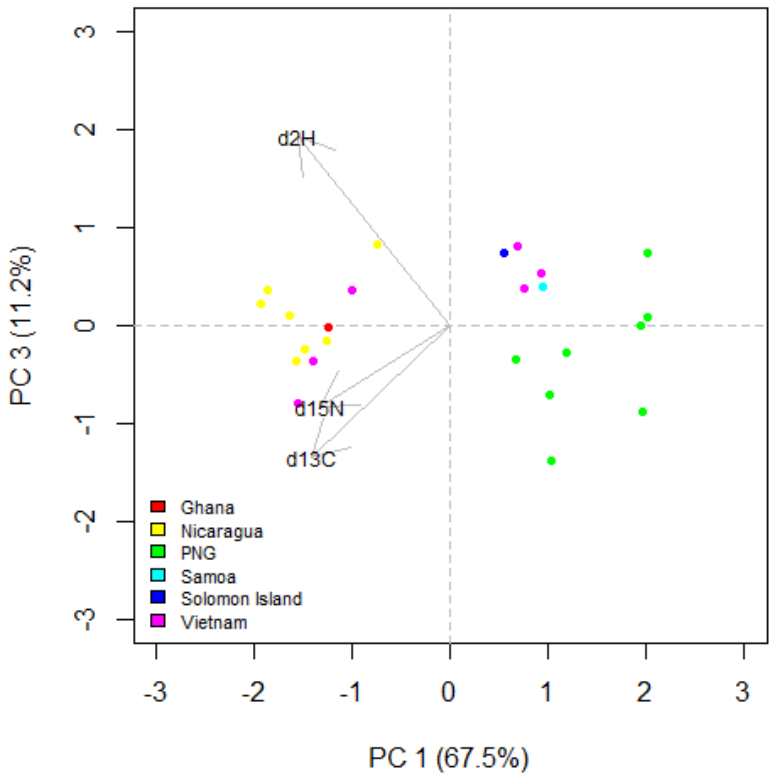


*Figure 4-5b. Biplot of PC1 vs PC3 for PCA on bulk $\delta^2H$, $\delta^{13}C$ and $\delta^{15}N$ of cocoa beans from Ghana, Nicaragua, PNG, Samoa, Solomon Islands and Vietnam.*

This chapter has shown that the stable isotope ratios of the whole cocoa bean does reflect the origin of the sample. It was also important to have an understanding of the variability at the different levels as especially seen for the samples from Vietnam, which has shown large range at all levels explored. When the PCA was applied it was then seen to split between the two main clusters. This is representation of stable isotopes showing geographical and climatic conditions of origin and not political boundaries.

In conclusion, 1. The variability of isotope ratios within a samples bag varied for each country. variability of significance was seen from $\delta^{13}C$ for Vietnam.

2. The variability within a country was seen not to be impacted by the variability within a bag excluding the variability from $\delta^{13}C$ for Vietnam, which may have been influenced by the high variance within a bag.

3. Variability was seen between countries indicating that isotope ratios reflect the origin.

4. The countries were seen to split into two groups. Ghana and Nicaragua and PNG, The Solomon Islands and Samoa. With samples from Vietnam split between the two.

# Chapter 5 – FAME results

Before stable isotope ratios of fatty acids can be explored as a way to determine the origin of cocoa beans in chocolate it is important to show that the stable isotope ratios of fatty acids in cocoa beans reflect their origin. As seen in the previous chapter, the isotope ratios of the whole unshelled cocoa bean were able to separate some countries of origins with others. The fatty acids Palmitic (C16), Stearic (C18) and Oleic (C18-1) are the three most abundant fatty acids in cocoa beans. These were extracted from the cocoa beans from Ghana, Nicaragua, PNG, The Solomon Islands, Samoa and Vietnam and then methylated. The isotope ratios $\delta^{2}H$ and $\delta^{13}C$ were measured on the methylated fatty acids using the GC-IRMS. The results from this can be found in appendix 2.

To explore the potential for discrimination based on origin, box and whiskers plots were created for each FAME for both $\delta^{2}H$ and $\delta^{13}C$ values. These graphs provided an understanding of what the variability is between countries as well as within countries for each FAME. These graphs indicated that there was some variability between the countries, to explore this variability further statistical analysis was applied to the results. Firstly, scatter plots of $\delta^{13}C$ verse $\delta^{2}H$ were created for each FAME to see if combining the isotope ratios provided a more robust separation between the countries. As these showed the potential for each FAME to separate out the samples based on origin, multivariate analysis in the way of PCA was then applied. The PCA showed clustering that enables some origins to be separated from other origins.

Aim: Investigate if stable isotope measurements of FAMEs reflect the country of origin for cocoa beans.

Objectives: Evaluate if stable isotopes of FAMEs from beans reflect origin.

1. Create box plots to explore the stable isotope ratios of H and C in the FAME
2. Combine the H and C results and explore any variability between countries
3. Conduct multivariate analysis to explore any clustering
4. Combine the FAME results with the Bulk cocoa bean results and evaluate clustering

## 5.1 FAME Hydrogen isotope ratios

To explore the variability between and within countries based on their $\delta^2H$ for the three individual FAME, box and whisker plots were created for each. Figures 5-1a, 5-1b and 5-1c display the range for C16, C18 and C18-1 respectively.

The $\delta^2H$ for C16 (Figure 5-1a) shows the countries split into two general groups; the first group containing Ghana, Nicaragua and Vietnam have higher $\delta^2H$ values than the second group containing PNG, Samoa and The Solomon Islands. The samples from Vietnam and PNG show greater variation within the country compared to Nicaragua. Both Samoa and The Solomon Islands had only one sample, and Ghana had three from same sample bag analysed. The results for the C18 (Figure 5-1b) show a similar picture with Ghana and Nicaragua having less negative values compare with PNG, Samoa and The Solomon Islands. Vietnam had values that fell between both groups. Again, PNG and Vietnam are seen to have greater variation within the country than Nicaragua. The range for PNG shows the upper and lower quartiles are sitting on the maximum and minimum values for this FAME. Vietnam also shows maximum and minimum values closer to the upper and lower quartiles compare with C16 and C18-1. The differences between countries in both C16 and C18 $\delta^2H$ values are not seen to the same extent for C18-1 (Figure 5-1c). However, Samoa and The Solomon Islands, which have similar values for both C16 and C18, have different values for C18-1. As only one sample was measured for these two countries, it is unknown if the difference is real or an artefact of the sample size, i.e. cannot tell from these data if they would also have large spread within the country as seen for Vietnam and PNG with this FAME.

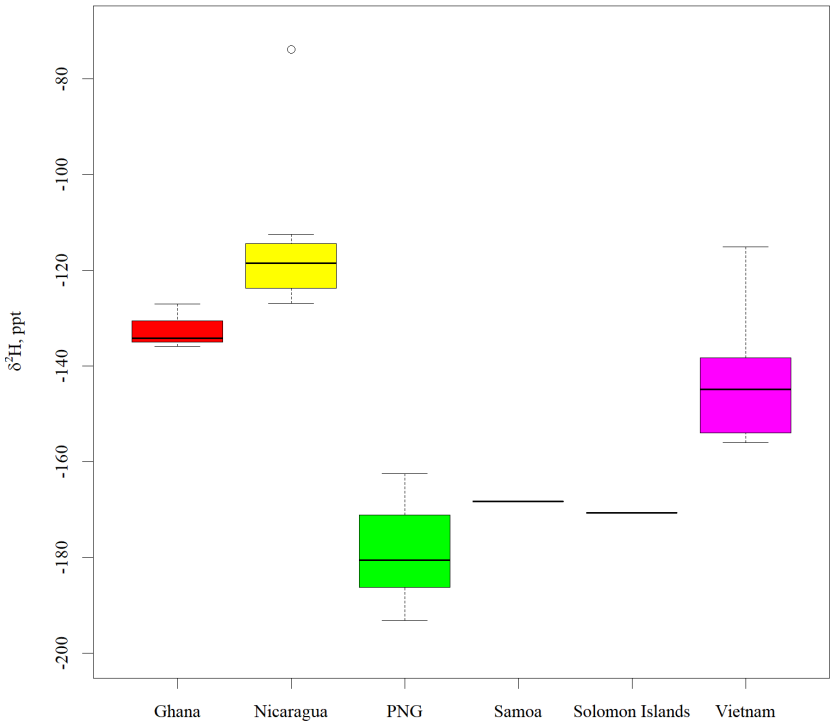


*Figure 5-1a Box and whisker plot of FAME C16 $\delta^2H$ values for Ghana(3), Nicaragua(6), PNG(6), Samoa(1), Solomon Islands(1) and Vietnam(5)*

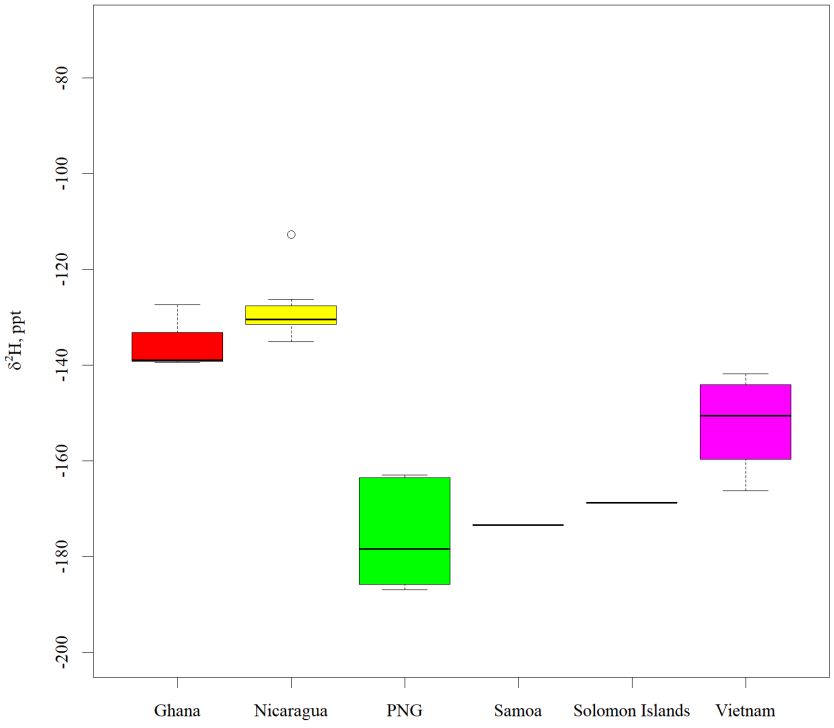


*Figure 5-1b Box and whisker plot of FAME C18 $\delta^2H$ values for Ghana(3), Nicaragua(6), PNG(6), Samoa(1), Solomon Islands(1) and Vietnam(5),*

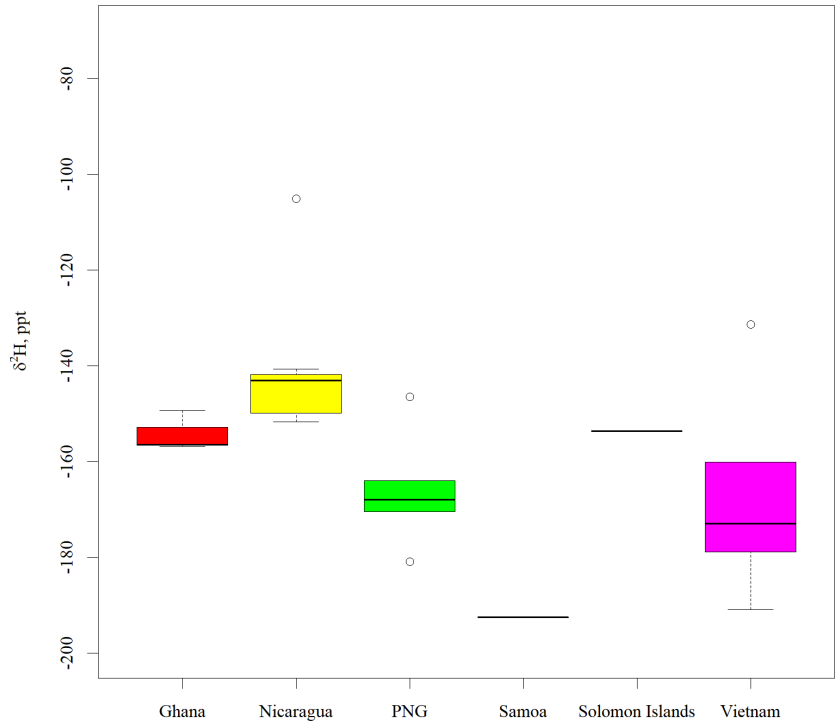


*Figure 5-1c Box and whisker plot of FAME C18-1 $\delta^2H$ values for Ghana(3), Nicaragua(6), PNG(6), Samoa(1), Solomon Islands(1) and Vietnam(5).*

The range seen for Vietnam covers the range of all the other countries for C18-1. Even though Nicaragua has a smaller range than PNG and Vietnam, a possible outlier can be seen for all FAMEs. A Grubbs test was performed on the samples from Nicaragua for each FAME and in each case gave a significant result that the sample was a possible outlier. The results of this sample do not fit within the same range as the other samples from Nicaragua. With the limited information provided on the origin of the sample, it cannot be said that this is not a valid sample. However, this sample was also positioned at the end of a run which had considerable instrumental drift on the later samples. Therefore, with the uncertainty in the reliability of this result for this current study, the sample was removed for any further analysis.

## 5.2 FAME Carbon isotope ratios

The variability for $\delta^{13}C$ of the three FAMEs were also explored, with the use of box and whisker plots. Figures 5-2a, 5-2b and 5-2c represent the $\delta^{13}C$ values for each country for C16, C18 and C18-1 respectively.

The $\delta^{13}C$ values for C16 in Figure 5-2a show that the sample from Samoa has a more negative value compared to the other countries. More samples would need to be measured to get an idea of the true range within the country. However, this sample shows a difference to the other countries which is greater than the spread for other countries such as PNG. The samples from Nicaragua have the least negative values from all countries, distinguishing from Ghana, PNG and The Solomon Islands. Vietnam has $\delta^{13}C$ values that overlap with the values from Nicaragua, Ghana and PNG; however, the mean $\delta^{13}C$ for Vietnam is more negative than that of Nicaragua. Nicaragua has the greatest range in $\delta^{13}C$ values. In comparison, Vietnam has a smaller range between its upper and lower quartile values. For C18 (Figure 5-2b) not a lot of variability is seen between countries, except for the sample from Samoa, which is again seen to have a much more negative $\delta^{13}C$ value than the other countries. The samples from Nicaragua and PNG have a similar range within their countries.

The $\delta^{13}C$ values for C18-1 (Figure 5-2c) divides the countries into two groups. Ghana, Nicaragua and Vietnam all have less negative $\delta^{13}C$ values with PNG, Samoa and The Solomon Islands having more negative $\delta^{13}C$ values. The samples from Vietnam, Nicaragua and PNG all appear to have similar range within the country.

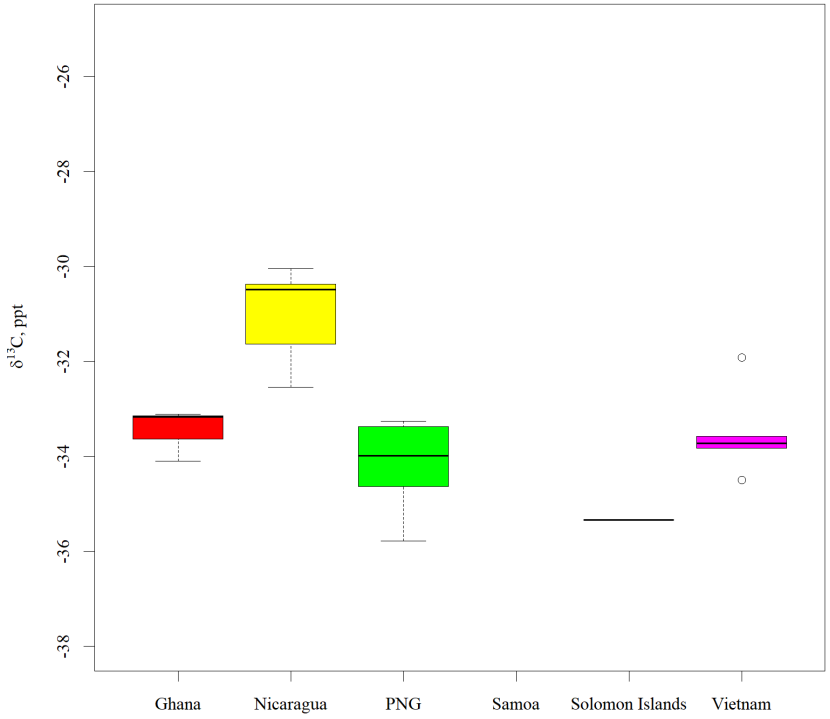


*Figure 5-2a Box and whisker plot of $\delta^{13}C$ values for FAME C16 extracted from cocoa beans of Ghana(3), Nicaragua(6), PNG(6), Samoa(1), Solomon Islands(1) and Vietnam(5).*

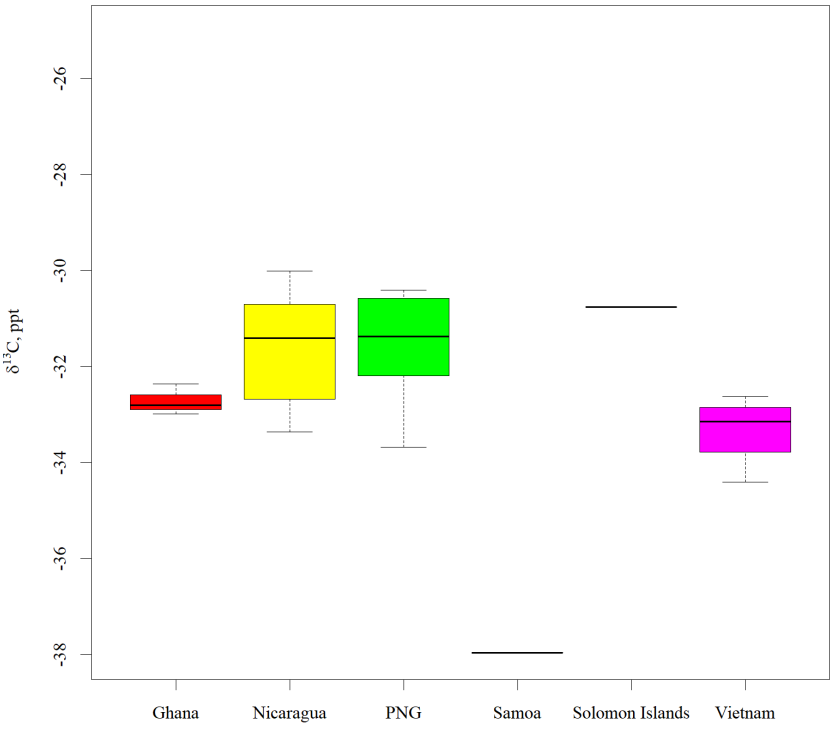


*Figure 5-2b Box and whisker plot of $\delta^{13}C$ values for FAME C18 extracted from cocoa beans of Ghana(3). Nicaragua(6), PNG(6), Samoa(1), Solomon Islands(1) and Vietnam(5).*

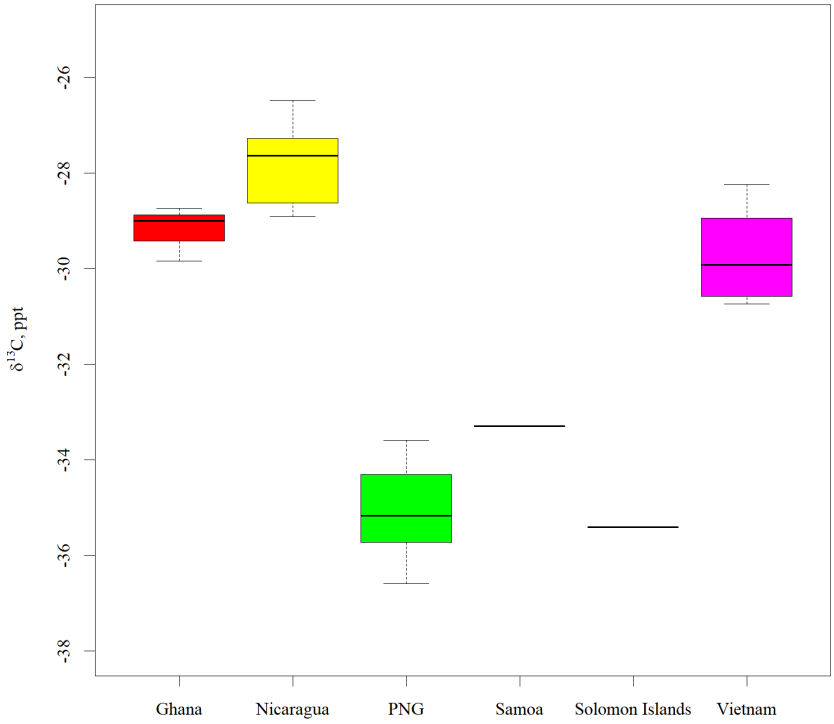


*Figure 5-2c Box and whisker plot of $\delta^{13}C$ values for FAME C18-1 extracted from cocoa beans of Ghana(3), Nicaragua(6), PNG(6), Samoa(1), Solomon Islands(1) and Vietnam(5).*

That Vietnam had relatively similar and lower range for the $\delta^{13}C$ values of all three FAMEs is important to note. When measuring the bulk values, there was a high variability for the $\delta^{13}C$ values across Vietnam. This indicates that the carbon that caused the variation within the bulk samples was not linked to the FAMEs within the cocoa samples. As analysing the FAMEs of the cocoa bean is a way that may allow for information on the cocoa bean within a chocolate mixture. It shows that these may not be affected by the sample issues as seen with the bulk sample.

## 5.3 Carbon Vs Hydrogen scatter plots

The initial exploration of the data using the box and whisker plots showed that there were differences between samples from different origins. This difference was further explored by combining the two isotope ratios for each FAME to see if this allowed for grouping of the samples based on origin. This analysis was done by creating scatter plots of $\delta^{13}C$ vs $\delta^{2}H$ for each FAME. Figure 5-3a, 5-3b and 5-3c are scatter plots for C16, C18 and C18-1 respectively.

Figure 5-3a shows the scatter plot of C16 values illustrating that the samples from each country are grouped. The samples from Nicaragua are seen to have less negative values for both the

$\delta^{13}C$ and $\delta^{2}H$. Where PNG and The Solomon Islands are seen to have the most negative $\delta^{2}H$ values. Samoa is seen to have the most negative $\delta^{13}C$ value. For C16, most of the separation occurs in the $\delta^{2}H$ values as most countries have $\delta^{13}C$ values between -33 and -35 ‰ or close to it. The scatter plot of C18 (Figure 5-3b) also shows separation due to the $\delta^{2}H$ values. The countries appear to split into two groups Nicaragua, Ghana and two samples from Vietnam in one group which have less negative $\delta^{2}H$ values. The other group contains PNG, The Solomon Islands, along with the rest of the samples from Vietnam; these samples have more negative $\delta^{2}H$ values. Samoa, which has a much more negative $\delta^{13}C$ than all other countries, is seen to be on its own. Unlike C16 and C18, the samples from C18-1 (Figure 5-3c) show separation based on their $\delta^{13}C$ values. The samples can be split into two groups. One group containing Nicaragua, Ghana and Vietnam which have less negative $\delta^{13}C$ values. The other group containing PNG, The Solomon Islands and Samoa are seen to have more negative $\delta^{13}C$ values. Most of the samples from Vietnam have more negative $\delta^{2}H$ values compared to Ghana and Nicaragua. Samoa is also seen to have a more negative $\delta^{2}H$ value.

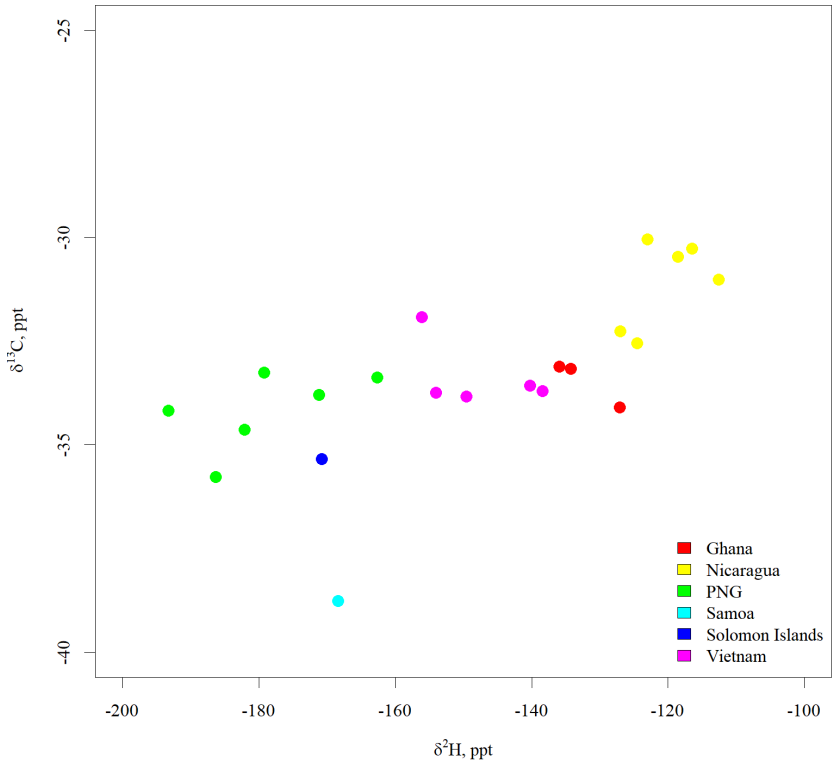


*Figure 5-3a Plot of $\delta^{13}C$ vs $\delta^{2}H$ values of FAME C16 extracted by Cocoa beans Ghana, Nicaragua, PNG, Samoa, Solomon Islands and Vietnam*

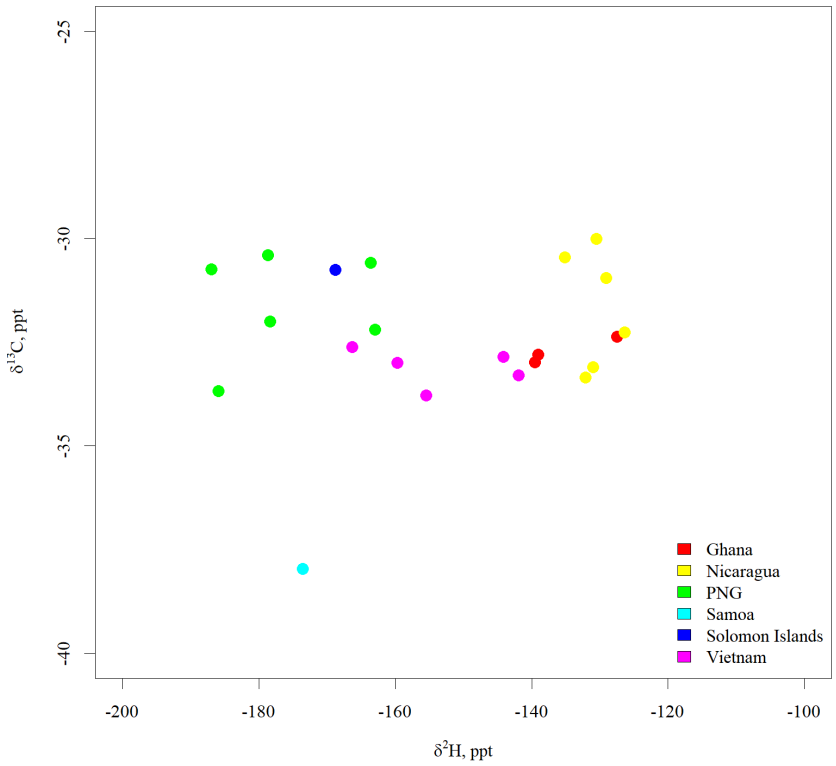


*Figure 5-3b Plot of $\delta^{13}C$ vs $\delta^{2}H$ values of FAME C18 extracted from Cocoa beans for Ghana, Nicaragua, PNG, Samoa, Solomon Islands and Vietnam*

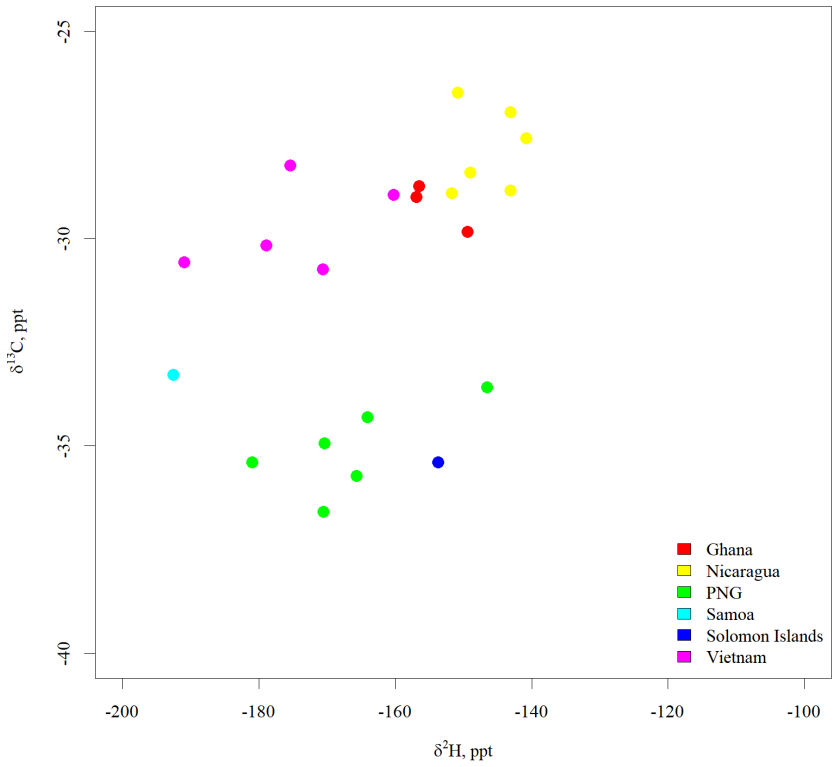


*Figure 5-3c Plot of $\delta^{13}C$ vs $\delta^{2}H$ values of FAME C18-1 extracted from Cocoa beans for Ghana, Nicaragua, PNG, Samoa, Solomon Islands and Veitnam*

All three FAMEs have shown similar trends in separating the countries from each other with Nicaragua, Ghana and Vietnam often in the same group and PNG and The Solomon Islands in a group. The sample from Samoa is out on its own in most of the scatter plots due to its more negative $\delta^{13}C$ value. Both sets of box and whisker plots and scatter plots show that the $\delta^{13}C$ and $\delta^{2}H$ values for the FAMEs can show some discrimination between the origins. However, just $\delta^{13}C$ or $\delta^{2}H$ was not able to show discrimination for all three FAMEs. C16 was able to show some variability between countries for both isotopes, where C18 showed variability in just the $\delta^{2}H$ and C18-1 showed variability for $\delta^{13}C$. When analysing the two isotopes together in the scatter plots Figures 5-3a, 5-3b and 5-3c, it can be seen that analysing multiple variables together allows for greater separation than achieved just analysing the one variable. Vietnam, for example, is seen to have around the same $\delta^{13}C$ values as Ghana and Nicaragua for C18-1 but more negative $\delta^{2}H$ value. Discrimination can be seen across the multiple parameters; therefore, a multivariate analysis was applied to this data.

## 5.4 PCA analysis

The previous analysis has shown the benefit of combining two variables to help improve the classification of the samples based on origin. To follow from this, a multivariate analysis was conducted to combine all the variables and explore if this achieves greater discrimination and classification of the samples based on origin. A PCA was conducted on the FAMEs $\delta^{13}C$ and $\delta^{2}H$ values. For this analysis, only the first two principal components (PC) were used; these explained 90% of the variance, as seen in Figure 5-4a. They were also the only PC to show a positive value when the log of the variances was graphed (Figure 5-4b). PC1 and PC2 were plotted in a biplot (Figure 5-5) to see if discrimination based on the origins of the cocoa beans could be achieved. Each sample is coloured according to the origin. Vectors of original variables, $\delta^{2}H$ for C16, C18 and C18-1 and $\delta^{13}C$ for C16, C18 and C18-1 are also included in the graph.

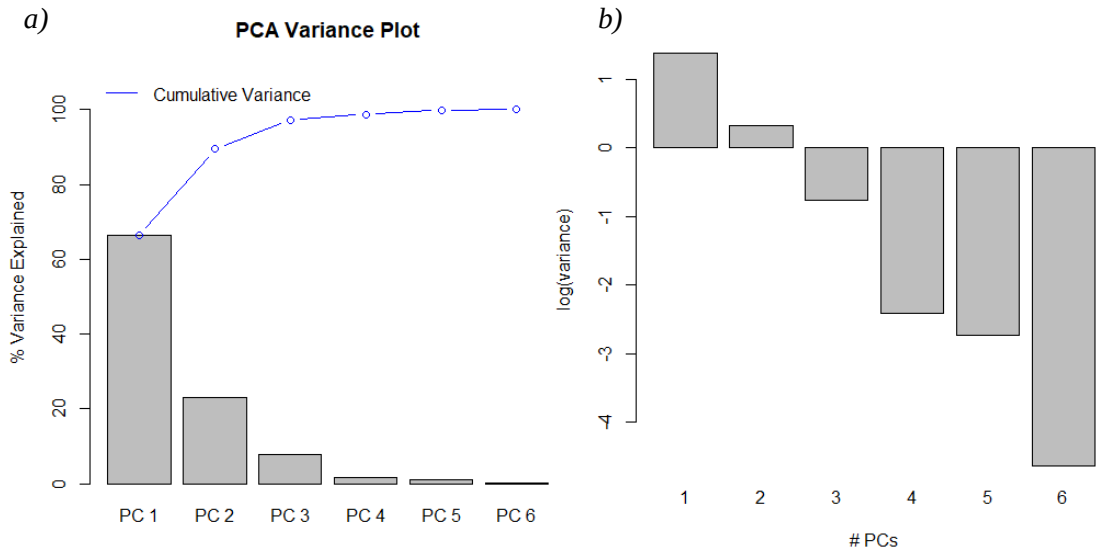


*Figure 5-4 a) Variance plot for PCA of Cocoa bean FAME $\delta^2H$ and $\delta^{13}C$ for all countries. b) Scree plot for PCA of cocoa bean FAMEs $\delta^2H$ and $\delta^{13}C$ for all countries*

Figure 5-5 shows that the samples from PNG and The Solomon Islands are grouped. They both show a positive PC1 and PC2 score. These samples are seen to be separated from the samples from Nicaragua and Ghana. The samples from Nicaragua have a negative PC1 score, where the PC2 score is seen to be zero for the centre of this group. The samples from Ghana are seen to have a negative PC1 and PC2 score. The samples from Vietnam have a negative PC2 score; the ellipse produced for the Vietnam samples is seen to overlap the ellipses for both Nicaragua and PNG. The samples from Ghana are seen to be sitting inside the ellipse for Vietnam. The only origin that is seen to have absolutely no overlap with any others is Samoa which is seen to have a positive PC1 score and negative PC2 score. The countries that had multiple samples analysed show very different spread between their samples. The samples from PNG are seen to have a relatively long and widespread, where the samples from Nicaragua follow long and narrow spread pattern. Vietnam is seen to have a smaller spread of samples.

The vectors in the graph show that the original variables can be split into three groups. All have a negative PC1 loading. $\delta^2H$ C18 and C16 along with the $\delta^{13}C$ C18-1 all show similar loadings with a negative PC2 loading and are close to perpendicular to $\delta^2H$ C18-1 and $\delta^{13}C$ C16 which have a positive PC2 loading. $\delta^{13}C$ C18 is also seen to have a positive PC2 loading however has a less negative PC1 loading than $\delta^2H$ C18-1 and $\delta^{13}C$ C16. It is seen that the loadings of $\delta^2H$ C18, C16 and $\delta^{13}C$ C18-1 are characteristic of the samples from Ghana.

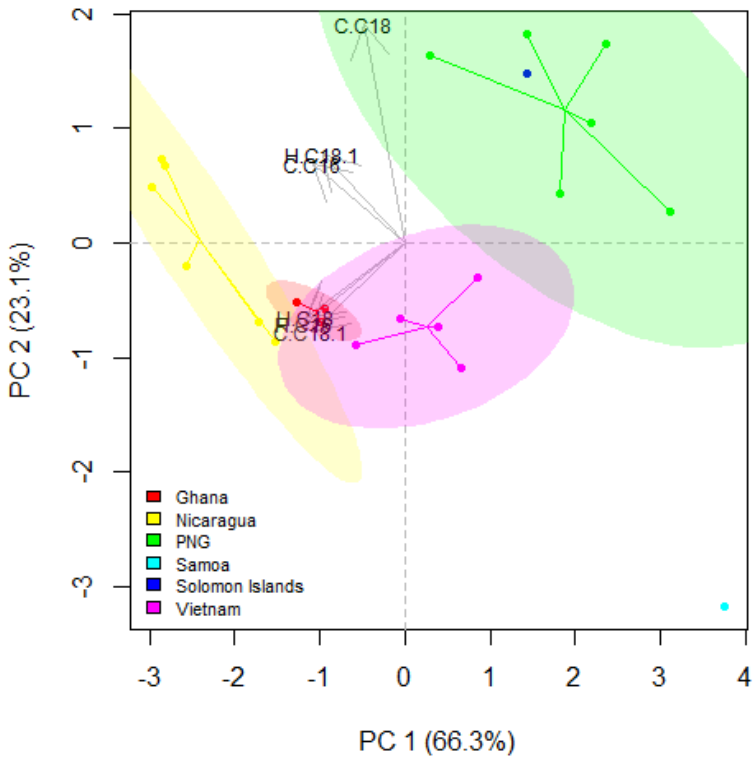


*Figure 5-5 Biplot of PC1 vs PC2 for PCA of Cocoa bean FAMEs $\delta^2H$ and $\delta^{13}C$ values for all countries.*

This PCA indicates that while not all origins can be separated from one another, there is some discrimination between countries based on their $\delta^{13}C$ and $\delta^2H$ values of the three FAMEs. The greatest being the separation between Nicaragua with PNG and The Solomon Islands and the sample from Samoa being wholly separated from all other origins. This study is limited with the number of samples analysed, especially for The Solomon Islands and Samoa, for which only one sample was analysed. However, it does show some indication on whether it is possible to discriminate the cocoa beans based on their origin using FAME isotope values.

## 5.5 Cocoa Beans: Combined Bulk and FAME analysis

To explore if the combination of the bulk and FAMEs analysis provided a stronger separation, a PCA was conducted that combined the bulk $\delta^2H$ and $\delta^{13}C$ results with the FAMEs result for the cocoa beans. As with the PCA for the FAME results, only the first two PCs were used. Combined these PC explain 85% of the variance as seen in Figure 5-6a. They also showed positive values on the Scree plot, where all others showed a negative value (Figure 5-6b).

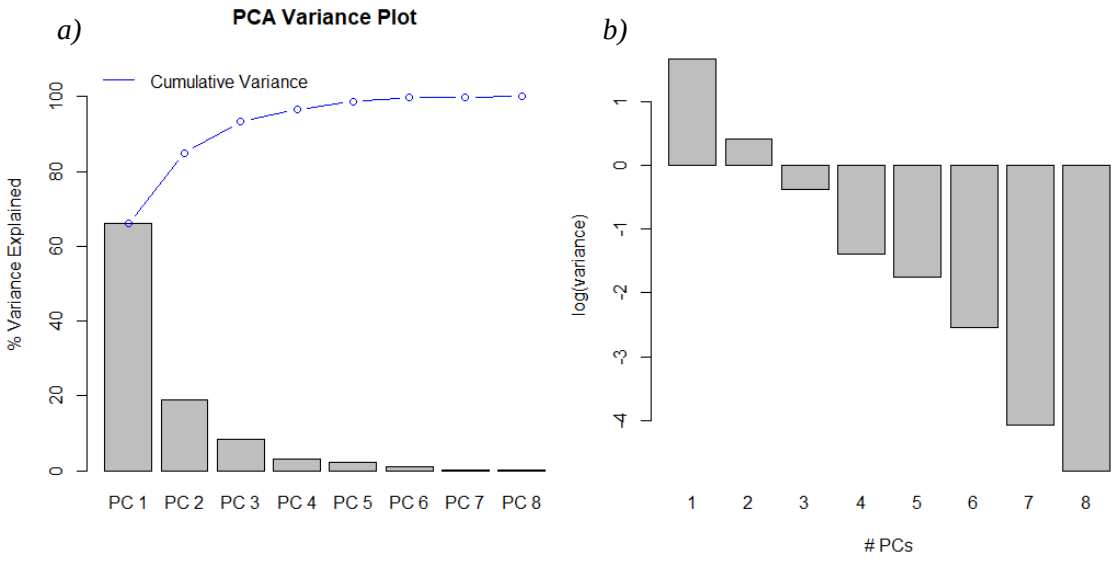


*Figure 5-6 a) Variance plot for PCA of Cocoa bean bulk and FAME $\delta^2H$ and $\delta^{13}C$ for all countries. b) Scree plot for PCA of cocoa bean bulk and FAME $\delta^2H$ and $\delta^{13}C$ for all countries*

With the addition of the bulk $\delta^2H$ and $\delta^{13}C$ results in FAME results, the PCA (Figure 5-7) still shows a similar picture. However, there is a greater separation between the samples from PNG with Vietnam. When just the FAMEs were analysed, these countries showed some overlap in the ellipses. Where the PCA for only the bulk $\delta$ values (Figures 4-5a and 4-5b) showed, Vietnam split across the two groups of countries. The sample from Ghana sits within the ellipse of Vietnam; this was also seen with just the FAMEs results (Figure 5-5). More samples from Ghana would need to be analysed to get a better understanding of the variation within this country and to get a better idea of how it is placed compared with Vietnam and Nicaragua. Again, as with the FAME PCA, the sample from Samoa is still out on its own.

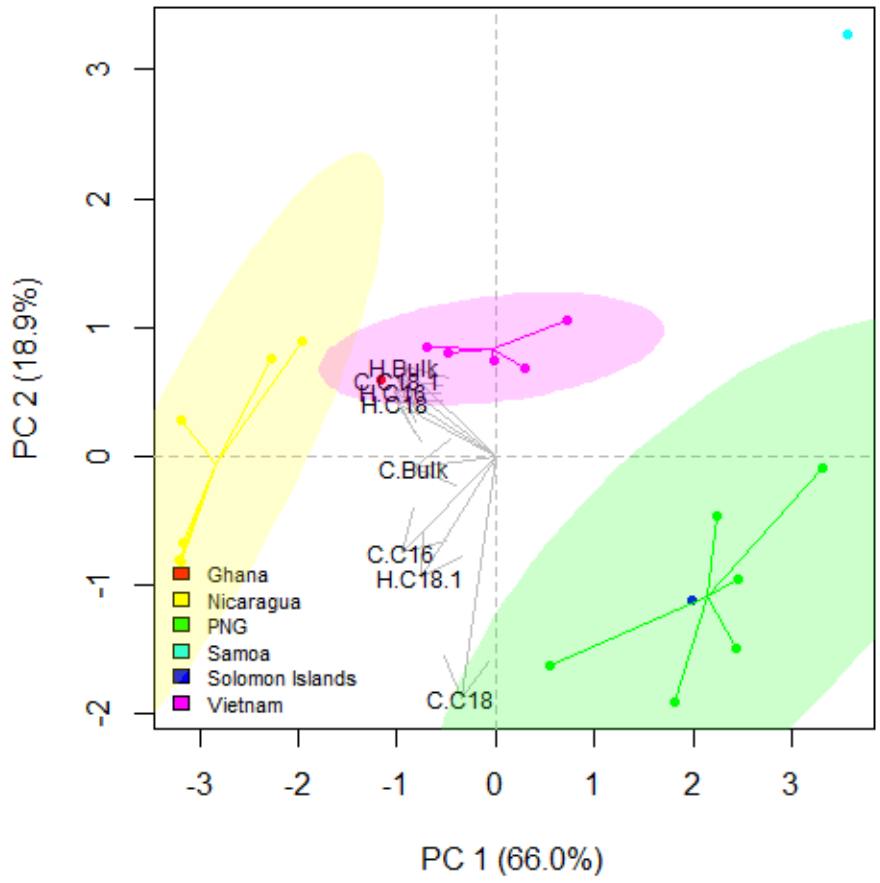


*Figure 5-7 Biplot of PC1 vs PC2 for PCA of cocoa bean bulk and FAME $\delta^2H$ and $\delta^{13}C$ values for all countries.*

The results of the FAME, bulk and combined analysis, enable the countries to be separated into two groups. PNG and The Solomon Islands in one group and Nicaragua, Vietnam and Ghana in the other. This separation shows that there is some variation between the cocoa beans based on origin with greater separation when more data is analysed together. The most significant limitation of this study is the small number of samples analysed with some countries only having one sample. Increasing the number of samples will allow for a greater understanding of the variability within the countries as well as between them. Isotopic variation is driven by geographic and environmental factors and not political borders. Thus, some countries could have different values across different regions. Therefore, analysing samples from different regions across a country is necessary for understanding the variability within the country. A much larger dataset that includes authenticated-origin samples from across the countries would enable separating samples based on regions instead of countries and hence allow for greater discrimination based on origin.

In conclusion, 1. The FAMEs C16 and C18 showed variation between countries for $\delta^2H$ and C16 and C18-1 showed variation between countries for $\delta^{13}C$

2. Combining the $\delta^2H$ and $\delta^{13}C$ values for each isotope provided separation between countries

3. The PCA of the FAME isotopes grouped the cocoa beans into two groups based on countries of origins. Nicaragua and Ghana in one group and PNG and The Solomon Islands in another. Vietnam was seen to be in between these two groups and Samoa out on its own.

4. Combining the FAMEs with the bulk results provided greater separation than separately with Vietnam seen to be closer to Nicaragua than PNG.

# Chapter 6 – Chocolate

Cocoa beans are not often sold to consumers as a bean; rather, they are sold as a manufactured product. One such product is chocolate, which is made up of cocoa beans and sugar as a minimum and will often have other components as well, such as milk. To investigate if the origin of cocoa beans is reflected in the fatty acids extracted from chocolate. The fatty acids of a couple of chocolate samples made in the laboratory were extracted and methylated. The previous chapter showed that it was possible to group chocolate samples based on origin. Therefore, the same technique was applied to the chocolate samples, and these results were then compared to the results from the cocoa bean FAME analysis.

The two chocolate samples were made in the laboratory using methods that mimicked those used by chocolate factories, for the full method used refer to chapter 3-Method section 3.3. The beans used were chosen at random from the bag of samples that originated from Ghana. These were then roasted and blended with icing sugar to make the chocolate samples. Icing sugar was used as the tapioca starch in the icing sugar created a third element to the mixture. One of the chocolate samples also had soy lecithin added, as this is often a common ingredient added to chocolate. The type of chocolate analysed is simple in that it only had a few ingredients compared to many found at the supermarket. However, it is a good proximation to test if the origin of the cocoa beans in the chocolate can be verified.

Aim: Investigate if stable isotope measurements can be used to verify the origin of cocoa extracted from Chocolate.

Objectives: Extract fatty acids from chocolate and evaluate if they reflect the origin of the cocoa beans.

1. Measure the $\delta^{2}H$ and $\delta^{13}C$ isotope ratios on the individual FAMEs
2. Test for the influence of lecithin
3. Explore the data to determine if the results cluster according to origin classification

Soy lecithin is often added during the process of making chocolate; therefore, it was essential to check that this ingredient was not going to affect the result in any way. Soy Lecithin is a mixture of phospholipids which meant that it might derivatise along with the fatty acids and be carried through the GC and analysed. The sample that had the soy lecithin added was found to have no extra peaks in the chromatogram. A sample of pure soy lecithin was processed and measured to check that the absence of an additional peak in the chromatogram was not due to the soy lecithin co-eluting one of the FAMEs from the cocoa. No peak in the chromatogram was produced from this ingredient.

The $\delta^{13}C$ and $\delta^{2}H$ were measured for the FAMEs in chocolate. These were then compared to the FAME results of the cocoa beans used to produce the chocolate. The $\delta^{13}C$ results of C16 showed the chocolate samples have a mean value closest to that of Ghana, followed by Vietnam (Table 6-1). For C18, the mean value was closest to that of Vietnam with Ghana being the next closest, and this is also seen for the C18-1 results. For $\delta^{2}H$ results, the chocolate samples are seen to be closest to Vietnam for all three FAMEs. Table 6-2 shows the standard deviation for both Ghana and Vietnam as well as the chocolate samples. It can be seen that the uncertainties for the chocolate samples are reliant on the instrument precision. For the $\delta^{2}H$ values, the standard deviation can be seen to be very high, which means that the instrument is less precise than that of the $\delta^{13}C$ values. It can also be seen that the uncertainties, in general, are much smaller for $\delta^{13}C$ values than $\delta^{2}H$. When looking at these uncertainties of these measurements, it shows that the chocolate samples could show values close to those of Ghana.

*Table 6-1. Mean of $\delta^{13}C$ and $\delta^{2}H$ values of FAME C16, C18 and C18-1 for cocoa beans of each country and sample of chocolate*

| Origin | Number of samples | $\delta^{13}C$ | | | $\delta^{2}H$ | | |
|---|---|---|---|---|---|---|---|
| | | C16 | C18 | C18-1 | C16 | C18 | C18-1 |
| **Chocolate** | 2 | -33.73 | -33.62 | -30.00 | -155.45 | -144.45 | -176.33 |
| **Ghana** | 3 | -33.46 | -32.72 | -29.20 | -132.40 | -135.30 | -154.22 |
| **Nicaragua** | 6 | -31.10 | -31.69 | -27.87 | -120.32 | -130.64 | -146.38 |
| **PNG** | 6 | -34.18 | -31.61 | -35.10 | -179.09 | -176.05 | -166.33 |
| **Samoa** | 1 | -38.77 | -37.97 | -33.29 | -168-34 | -173.491 | -192.47 |
| **Solomon Islands** | 1 | -35.34 | -30.75 | -35.41 | -170.73 | -168.73 | -153.71 |
| **Vietnam** | 5 | -33.36 | -33.12 | -29.74 | -147.63 | -153.48 | -175.19 |

*Table 6-2. Standard deviation of $\delta^{13}C$ and $\delta^{2}H$ values of FAME C16, C18 and C18-1 for cocoa bean samples from Ghana and Vietnam*

| Origin | $\delta^{13}C$ standard deviation | | | $\delta^{2}H$ standard deviation | | |
|---|---|---|---|---|---|---|
| | C16 | C18 | C18-1 | C16 | C18 | C18-1 |
| **Chocolate** | 0.40* | 0.29* | 0.29 | 10.70* | 9.52* | 20.93* |
| **Ghana** | 0.75* | 0.61* | 0.15 | 2.99 | 6.27 | 9.51* |
| **Vietnam** | 0.29 | 0.61* | 0.94 | 8.99 | 9.41 | 18.2 |

*Precision value used to calculate standard deviation as variance was less than the analytical precision value; therefore, population variance irrelevant.

The data from the chocolate samples were put through a blind test where they were added to the PCA model created for the FAMEs earlier Figure 6-1. From this test, it was seen that the samples were still unable to be distinguished from that of Vietnam. The samples from Ghana were also found to be indistinguishable from that of the Vietnam samples. However, the Chocolate samples had a positive PC1 score where the samples from Ghana had a negative PC1 score.

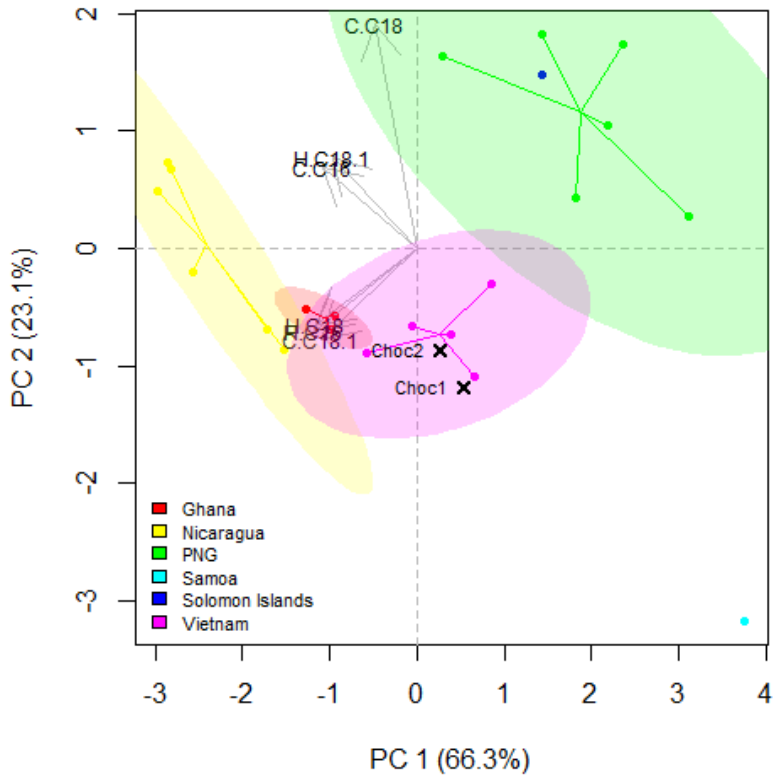


*Figure 6-1 Biplot PC1 vs PC2 from PCA of FAME isotope values extracted from cocoa beans with FAME isotope values from chocolate samples added indicated with an* ×

Even though the chocolate samples are not grouped with Ghana as the country of origin, they are still within a cluster that included Ghana. As only a small number of samples were analysed from Ghana, it limits the knowledge of the actual range within the country. Also, for the cocoa bean analysis, each sample is a single bean where the chocolate was made from multiple beans that had been roasted and blended to a homogenous liquid. This process might cause some fractionation of the isotopes if there was partitioning or material losses. More cocoa beans and chocolate samples from Ghana as well as other countries would need to be analysed and compared to determine if any measurable fractionation is occurring. If there is there may also be a pattern to this fractionation that an adjustment could be calculated to help with determining the origin based on a database of cocoa bean δ values.

In conclusion 1. The fatty acids were extracted from chocolate made of a known origin and $\delta^2H$, and $\delta^{13}C$ were measured on the FAMEs

2. Lecithin was found not to appear on the chromatogram and therefore found not to influence the results.

3. The chocolate samples made from cocoa beans originating in Ghana, were found to cluster with samples from Vietnam. The samples from Ghana were also found to cluster with Vietnam.

# Chapter 7 – Underivatised Free Fatty Acids

The fatty acid isotope results on the cocoa bean as seen previously in chapter 5 have shown some differences between the cocoa beans based on origin. However, this method involved the samples undergoing methylation to form methyl esters as these are volatile and allow for separation through the GC column. This process adds one carbon and three additional hydrogen atoms to the fatty acid. The isotope results provided from by the methyl esters include this additional methyl group. Hence mass balance calculations are required to remove the contribution related to methyl group to those from the cocoa beans to obtain the value of just the cocoa bean fatty acid. This process also requires extra preparation that can add to the amount of human error that could occur. The ideal situation would be to measure the isotope values without having to methylate the samples first. To test this, fatty acid samples were extracted and measured without derivatisation using a fatty-acid specific GC column (CP-FFAP).

Aim: Investigate the analysis of cocoa bean fatty acids without the need to derivatise them

Objectives: Extract and measure the FA from cocoa beans without derivatisation to see if

1. Usable chromatography from underivatized samples is obtained
2. Sufficient FA can be obtained from cocoa beans for measurements
3. Explore the δ values, ensure they make sense
4. Compare FA results with FAME results

The chromatograph (Figure 7-1) showed a couple of peaks with peak one (*1850.9 s)* at low concentration and peak two (*2048.0 s)* about the same as the C19 standard (*2110.34 s*). The peaks here are likely to show only the free fatty acids from the cocoa bean instead of the full amount. The rest of the fatty acids are likely to occur as triglyceride combinations. The methylation process used must also break down these triglyceride compounds to form the methyl esters. Cocoa beans have three major fatty acids, with only two peaks seen there could be the possibility that two of the fatty acids have the same retention time, or that one is more likely to always occur as a triglyceride than a free fatty acid.

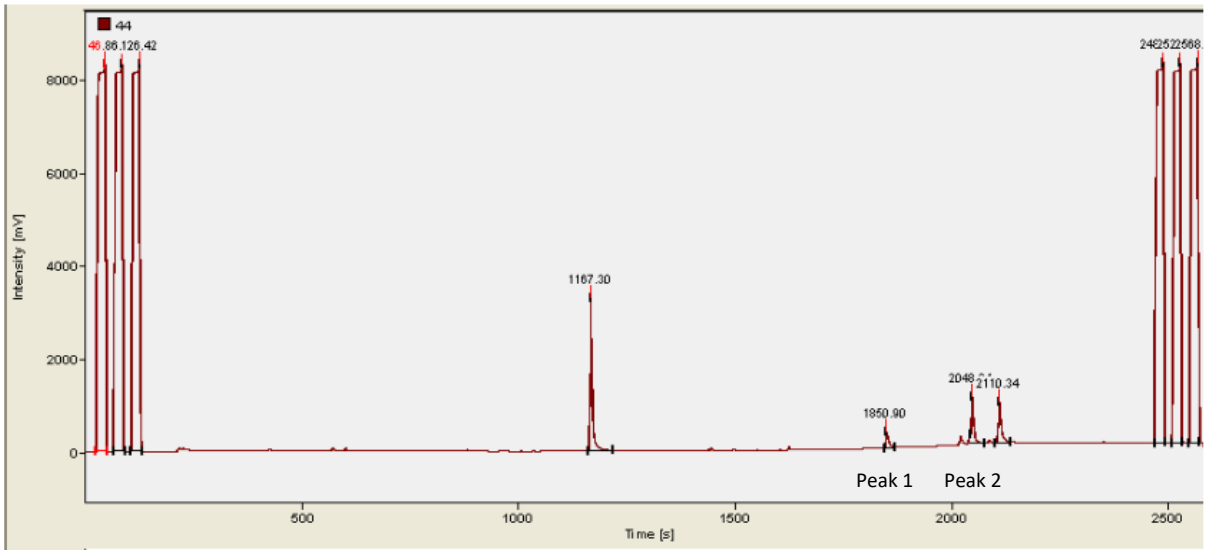


*Figure 7-1 Chromatograph of cocoa bean fatty acid sample from PNG. With two sample peaks shown Peak 1 – 1850.9 and Peak 2 – 2048.0. All other peaks are reference and internal standards.*

The two peaks seen were peak 1, retention time 1850.9 and peak 2 retention time 2048. However, not all samples provided detectable peaks, only a few samples had detectable peaks for peak 1, where most did for peak 2, as shown in Table 7-1. The sample from Samoa showed no detectable peaks. Figure 7-2 shows the $\delta^{13}C$ values of peak 2 for the underivatised cocoa beans. This graph shows that Nicaragua and Vietnam have less negative values, whereas PNG, Ghana and The Solomon Islands have more negative values. Only two samples from Ghana were analysed; however, Ghana shows a greater range than Nicaragua and Vietnam, which had more samples analysed. Ghana has also been shown to have values close to that of Nicaragua for bulk and derivatised samples, where these results show it having values closer to PNG and The Solomon Islands. PNG also shows a relatively high range compared with Nicaragua and Vietnam.

*Table 7-1: Number of samples with detectable peaks and mean $\delta^{13}C$ values of fatty acids extracted from cocoa beans.*

| Origin | Peak 1 | | Peak 2 | |
|---|---|---|---|---|
| | No. Detectable samples | $\delta^{13}C$ mean | No. Detectable samples | $\delta^{13}C$ mean |
| Ghana | 2 | -35.70 | 2 | -41.53 |
| Nicaragua | 2 | -35.46 | 6 | -32.71 |
| PNG | 2 | -34.83 | 7 | -39.28 |
| Samoa | - | - | - | - |
| Solomon Islands | - | - | 1 | -43.45 |
| Vietnam | 3 | -36.47 | 6 | -35.41 |

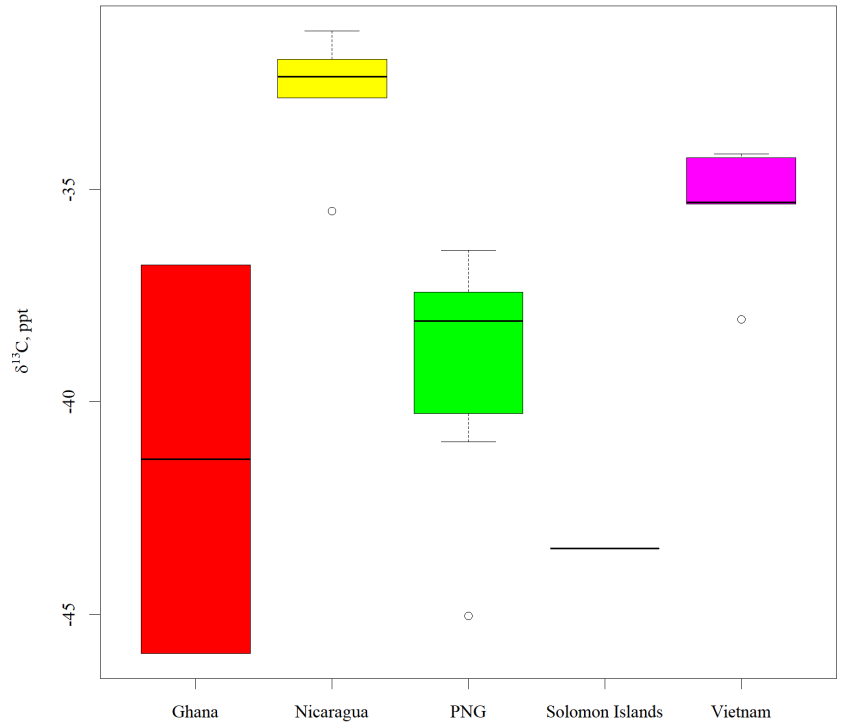


*Figure 7-2 Box and whisker plot of $\delta^{13}C$ values for peak 2 of fatty acid extracted from cocoa bean samples. Ghana(2), Nicaragua(6), PNG(7), Solomon Islands(1) and Vietnam(6)*

As not all fatty acids in the cocoa beans are analysed in this method fractionations between the free fatty acids and the fatty acids in the triglycerides may occur. This would limit the information that can be provided. Even so, it can be seen that some separation between the origins can be seen for the fatty acid delta values. Some improvements are needed for this technique to be fully effective. The amount of sample extracted needs to be increased to get

more detectable peaks for each sample, also improving method to ensure no fractionation occurs during the extraction phase. This method only shows the δ values of the free fatty acids, which means the majority of the fatty acids are not measured. Developing a method that could break down the triglycerides without the fatty acids forming other molecules like methyl esters would allow for the use of this column and the measuring of the fatty acids, without having to manipulate the results to remove δ values additional elements.

The separation seen of the FA samples shows similar patterns to that seen for both bulk $\delta^{13}C$ and the FAME results. Nicaragua is seen to have the least negative results, Vietnam is also seen to have less negative results compared with PNG and The Solomon Islands. This agrees with findings from the bulk and FAME results apart from results from C18 where minimal variability between countries was seen. The samples from Ghana show a different pattern to those seen in the other results. However, these results also show a substantial range relative to the other countries and only two samples were analysed for Ghana.

The aim of chapter was to investigate the analysis of cocoa beans fatty acids without the need to derivatise them. Therefore, removing the need to conduct mass balance calculations and the uncertainty that comes with it. The chromatogram showed a couple of peaks were obtained with sufficient amount for most samples for one of these peaks to be analysed. The results of this peck showed similar pattern to that of $\delta^{13}C$ of both bulk and FAME results. Therefore, this technique has shown some promise that with some more investigating and improvements, it may be able to be used as a means to determine the origin of cocoa beans.

In conclusion, 1. The chromatogram produced from the underivatized samples showed two peaks.

2. Of the two peaks formed one peak had sufficient amount of FA from most samples to be analysed. The other only some samples showed sufficient amount of FA.

3.The δ values for most countries provided reasonable results. Excluding Ghana, which was seen to have very high range for only two samples being analysed.

4. Most countries showed similar pattern of results as seen previously for the Bulk and derivatised $\delta^{13}C$ results.

# Chapter 8 – Conclusion and Future Work

## 8.1 Conclusion

The aim of the current study was to explore if it was possible to determine the origin of cocoa beans within a bar of chocolate using isotope ratios. To achieve this aim three objectives were identified: 1) Determine if stable isotope ratios do reflect the geographical origin of cocoa beans, 2) Evaluate if stable isotopes of fatty acids from beans reflect origin., and 3) Extract fatty acids from chocolate and evaluate if they reflect the origin of the cocoa beans.

To see if isotope ratios of cocoa beans do reflect the geographical origin of said cocoa beans. $\delta^{2}H$, $\delta^{13}C$ and $\delta^{15}N$ were measured on the cocoa nib of cocoa bean samples from Ghana, Nicaragua, PNG, Samoa, The Solomon Islands and Vietnam. The results of the bulk analysis showed that there were some differences between countries with them mostly splitting into two groups Nicaragua, Ghana and some samples from Vietnam in one group and PNG, Samoa, The Solomon Islands and the rest of the samples from Vietnam in the other. These two groups likely reflect the geographical and climatic factors of the countries. PNG, The Solomon Islands and Samoa are all island nations in the Pacific Ocean. Whereas Ghana, Nicaragua and Vietnam are all countries within a larger continental setting. With Vietnam being a long coastal country with a coast along the Pacific Ocean. The distinction observed for samples from within Vietnam may be due to the cocoa beans coming from different regions within the country. Isotope ratios reflect the geographical, not political boundaries; thus, different isotope values may occur within the same country if there is sufficient range in climatic and geographical conditions. There were relatively large spreads seen for both $\delta^{13}C$ and $\delta^{15}N$ for Vietnam. Geographical and climatic conditions are not the only thing that can influence the isotope ratios. They can also be affected by human impacts; the sample from The Solomon Islands is seen to have a much more positive $\delta^{15}N$ compared to that of PNG and Samoa. $\delta^{15}N$ can be highly influenced by agricultural practises, with the type of fertilizer having a significant effect on the values. The samples from Vietnam also show possible impact of human interference. High range in $\delta^{13}C$ was seen within a single bag, this variability may have been due to the fermentation process of the beans. Three beans from the one bag were analysed at random. The beans were seen to be physically different colours; the brown colour, which is often associated with chocolate is of a fermented cocoa bean, where one of these cocoa beans was more purple in colour, which is associated with a bean that had undergone less fermentation.

To analyse the stable isotope fatty acids in cocoa beans to determine if they reflect the origin, three fatty acids C16, C18 and C18-1, which are the three most abundant fatty acids in cocoa beans, were extracted from the cocoa beans and derivatised to form the methyl esters. Both the $\delta^{13}C$ and $\delta^{2}H$ were measured and analysed for all three fatty acids via compound-specific isotope ratio mass spectrometry. The results of this analysis showed that C16 and C18 showed some separation between countries for $\delta^{2}H$ and C16, and C18-1 showed separation between countries for $\delta^{13}C$. C18-1 for $\delta^{2}H$ and C18 for $\delta^{13}C$ show less variability between the countries. Even though Vietnam was seen to have a large range within the country for $\delta^{13}C$ in the bulk analysis, this is not seen for the fatty acid analysis, which means that the factors that influenced the variability for the bulk do not affect the fatty acids. From the PCA two groups were formed, Nicaragua and Ghana formed one group and PNG, and The Solomon Islands formed the other. The samples from Vietnam fell between the two groups, and the sample from Samoa appears out on its own. When the bulk $\delta^{2}H$ and $\delta^{13}C$ were added to the PCA, this increased the separation between PNG and Vietnam. FAMEs were used in this study due to their volatile nature and allow for separation through the GC column. However, this process adds a methyl group to the fatty acid, which requires mass balance calculations to remove the additional elements from the results. Fatty acids were run on a particular GC column CP-FFAP for $\delta^{13}C$. The results of this run showed lower concentrations than was seen for the FAME as it only accessed the free fatty acids from the cocoa bean. However, the results of one of the peaks showed some potential for separating the samples with the samples from Nicaragua and Vietnam, showing less negative values than the other countries.

The fatty acids and bulk isotope ratios of the cocoa bean can separate cocoa beans based on their origin. Chocolate samples were made from cocoa beans of known origin and the fatty acids were extracted and analysed against the cocoa bean results. Two chocolate samples were made in the lab by roasting and then blending cocoa beans from Ghana with icing sugar. One of the chocolate samples also included Lecithin, which is often used in commercialised chocolate as an emulsifier. The fatty acids were extracted and derivatised then $\delta^{2}H$ and $\delta^{13}C$ were measured. The results showed that the chocolate samples had values that were closest to that of Vietnam. However, only a small number of samples were measured for Ghana, and these too were found to be close in value to that of Vietnam. Therefore, more samples of beans and more samples of chocolate with known origins were analysed. Then it would provide a better indication as to whether this difference was due to the cocoa beans used or if a calculable

fractionation occurs during the process of chocolate making. The sample with lecithin showed no extra peaks from the addition. Meaning that if samples have this added, it will not show up on the GC chromatogram but also that it will not impact the results of the chocolate samples.

In conclusion, 1) Measurements of the bulk stable isotopes were able to differentiate cocoa beans according to their geographical origin.

2) Measurements of the fatty acid isotope ratios of cocoa beans were able to differentiate cocoa beans according to geographical origin

3) Fatty acids were extracted from chocolate and indicated the potential for identifying the origin of cocoa beans.

Therefore, this study fulfilled the objectives and achieved the aim of demonstrating that fatty acid stable isotope ratios are useful for differentiation of cocoa beans by the origin and that the measurement of the ratio of the stable isotopes of fatty acids extracted from chocolate show promise for determining the origin of the bean used to make the chocolate.

## 8.2 Limitations

This study has shown potential in using isotope ratios of both bulk and individual fatty acids as a way to determine the origin of cocoa beans. However, there was one major limitation to this study; the number of samples. There was only a small number of samples analysed for each country with some countries only having one sample analysed. The lack of samples limits the statistical robustness of the results and the conclusions made from them. For example, the sample from Samoa was seen to have much more negative $\delta^{13}C$ for the fatty acids C16 and C18. Without analysing more samples from Samoa, it is unclear whether this difference is for the country or just this one sample. Having more samples would also improve on understanding what the chocolate results are saying. Currently, they appear to be closer to the result of Vietnam, but if more samples of both chocolate and cocoa beans were analysed, then it would be easier to see if the results do reflect those of Ghana or not and if not is it due to fractionation occurring during the production of chocolate.

Another limitation was the limited information provided about the origins of each sample. The country of origin was provided for each sample, but not a lot of other useful information was

provided for most samples. The more information provided; the more detailed analysis of the results can be. If the exact location of each sample was known, GPS coordinates being the best method; then it could provide a better idea of possible variability from different regions within a country as isotope ratios show geographical boundaries, not political. The exact location is not the only information that would provide more insight, knowing the year and season of harvest can also be helpful. It would allow for information about climate during the growth of the beans to be obtained. If the samples are from different harvesting seasons, it could mean that the variability seen is more extensive than if they were from the same season.

### 8.3 Future work

With the potential seen in this current study, future studies would need to look at analysing more samples and eventually creating a database of bulk and fatty acid ratios of cocoa beans. To implement a system for auditing origin, it is imperative that the database referred to consists of samples of authentic origin, preferably to the farm of origin. A large enough database of authentic-origin samples would allow for possible traceability of samples with unknown origin information. Compiling an extensive database of many hundreds to thousands of samples would likely be too much work for one laboratory, and so collaboration between laboratories would be required. For this to work requires very careful consideration of quality control (QC) issues to ensure reproducibility and accuracy within and between laboratories. Specific quality control standards would need to be created, characterised and used with all samples. Of specific importance is the need for equilibration experiments for characterisation of exchangeable H for $\delta^2H$ of bulk samples.

The current study analysed individual beans from within sample bags. This method has its limits as it does not always provide the best method to ensure the whole population is represented. Another method that could provide additional information is by grinding up a large number of beans from the same sample bag together then subsample from the ground mass. The latter method was not used in the current study due to the technology not being available at the time. The composition of cocoa beans means that when it is being ground down the heat and pressure causes the fatty acids to melt and bind the ground cocoa beans back together. Which would make it harder to ensure that the subsample was homogeneous with the rest of the ground

samples. With technology improving every day future studies on cocoa beans could look at the latter method of subsampling

In the current study, the underivatised samples of cocoa beans FA were explored to see if they could be used instead of the methyl esters. The results show that there was some potential that the fatty acids could be separated without derivatisation. However, more work would need to be completed on this in future for this to be able to be used as a traceability tool. The results only showed the free fatty acids which could be influenced by fractionation effects within the cocoa bean. Finding a way to break up the triglycerides without, adding additional elements could allow for more accurate and precise results to the FAMEs studied.

Chocolate can be made up of many additional ingredients to cocoa beans and sugar. One common ingredient is milk. Milk also contains fatty acids which could lead to fatty acid results showing up both the milk and cocoa butter fatty acids. Therefore, future work needs to analyse the effects of adding milk to chocolate when analysing the fatty acids of cocoa beans. On top of the addition of milk, chocolate can be made using different processes. Some manufactures separate the cocoa powder from the cocoa butter before making the final product, which could lead to cocoa butter coming from different supplier to the cocoa powder. Therefore, this current technique will only work on identifying the origins of the cocoa butter in the chocolate. Looking at other compounds within the chocolate like the protein may provide a way also to identify the origin of the cocoa powder.

# References


Alexandre, P. (2020). Stable Isotopes and the Hydrosphere. In P. Alexandre (Ed.), *Isotopes and the Natural Environment* (pp. 21–28). Springer International Publishing. https://doi.org/10.1007/978-3-030-33652-3_3

Amankwah-Amoah, J., Debrah, Y. A., & Nuertey, D. (2018). Institutional Legitimacy, Cross-Border Trade and Institutional Voids: Insights from the Cocoa Industry in Ghana. *Journal of Rural Studies*, *58*, 136–145. https://doi.org/10.1016/j.jrurstud.2018.01.002

Amundson, R., Austin, A. T., Schuur, E. a. G., Yoo, K., Matzek, V., Kendall, C., Uebersax, A., Brenner, D., & Baisden, W. T. (2003). Global patterns of the isotopic composition of soil and plant nitrogen. *Global Biogeochemical Cycles*, *17*(1). https://doi.org/10.1029/2002GB001903

*Beans To Bar*. (n.d.). Retrieved May 20, 2020, from https://www.whittakers.co.nz/en_NZ/beans-to-bar/

Beckett, S. T. (2000). *The science of chocolate*. RSC.

Benson, S., Lennard, C., Maynard, P., & Roux, C. (2006). Forensic applications of isotope ratio mass spectrometry—A review. *Forensic Science International*, *157*(1), 1–22. https://doi.org/10.1016/j.forsciint.2005.03.012

Bertoldi, D., Barbero, A., Camin, F., Caligiani, A., & Larcher, R. (2016). Multielemental fingerprinting and geographic traceability of Theobroma cacao beans and cocoa products. *Food Control*, *65*, 46–53. https://doi.org/10.1016/j.foodcont.2016.01.013

Bligh, E. G., & Dyer, W. J. (1959). A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Canadian Journal of Biochemistry and Physiology*, *37*(1), 911–917. https://doi.org/10.1139/y59-099

Bontempo, L., Paolini, M., Franceschi, P., Ziller, L., García-González, D. L., & Camin, F. (2019). Characterisation and attempted differentiation of European and extra-European olive oils using stable isotope ratio analysis. *Food Chemistry*, *276*, 782–789. https://doi.org/10.1016/j.foodchem.2018.10.077

Caligiani, A., Palla, L., Acquotti, D., Marseglia, A., & Palla, G. (2014). Application of 1H NMR for the characterisation of cocoa beans of different geographical origins and fermentation levels. *Food Chemistry*, *157*, 94–99. https://doi.org/10.1016/j.foodchem.2014.01.116

Cambrai, A., Marcic, C., Morville, S., Sae Houer, P., Bindler, F., & Marchioni, E. (2010). Differentiation of Chocolates According to the Cocoa's Geographical Origin Using Chemometrics. *Journal of Agricultural and Food Chemistry*, *58*(3), 1478–1483. https://doi.org/10.1021/jf903471e

Carter, J., & Barwick, V. (2012). *Good Practice Guide For Isotope Ratio Mass Spectrometry*.

Carter, J. F., & Chesson, L. A. (2017). *Food Forensics: Stable Isotopes as a Guide to Authenticity and Origin*. CRC Press. https://doi.org/10.1201/9781315151649

Chaiseri, S., & Dimick, P. S. (1989). Lipid and hardness characteristics of cocoa butters from different geographic regions. *Journal of the American Oil Chemists' Society*, *66*(12), 1771–1776. https://doi.org/10.1007/BF02660745

Cidell, J. L., & Alberts, H. C. (2006). Constructing quality: The multinational histories of chocolate. *Geoforum*, *37*(6), 999–1007. https://doi.org/10.1016/j.geoforum.2006.02.006

*Cocoa Life—Cocoa Life is Empowering Women*. (n.d.). Cocoa Life. Retrieved May 25, 2020, from http://www.cocoalife.org/the-program/womens-empowerment

*Cocoa Life—Why Cocoa Life?* (n.d.). Cocoa Life. Retrieved May 25, 2020, from http://www.cocoalife.org/

Criss, R. E. (1999). *Principles of stable isotope distribution*. Oxford University Press.

Diakite, S. (2017). Does the UTZ label guarantee a child labor free product? *UTZ*. https://utz.org/better-business-hub/marketing-sustainable-products/does-the-utz-label-guarantee-a-child-labor-free-product/

Diomande, D., Antheaume, I., Leroux, M., Lalande, J., Balayssac, S., Remaud, G. S., & Tea, I. (2015). Multi-element, multi-compound isotope profiling as a means to distinguish the geographical and varietal origin of fermented cocoa (Theobroma cacao L.) beans. *Food Chemistry*, *188*, 576–582. https://doi.org/10.1016/j.foodchem.2015.05.040

Ehleringer, J. R., & Dawson, T. E. (1992). Water uptake by plants: Perspectives from stable isotope composition. *Plant, Cell & Environment*, *15*(9), 1073–1082. https://doi.org/10.1111/j.1365-3040.1992.tb01657.x

Ehtesham, E., Hayman, A. R., McComb, K. A., Van Hale, R., & Frew, R. D. (2013). Correlation of Geographical Location with Stable Isotope Values of Hydrogen and Carbon of Fatty Acids from New Zealand Milk and Bulk Milk Powder. *Journal of Agricultural and Food Chemistry*, *61*(37), 8914–8923. https://doi.org/10.1021/jf4024883

Elliott, P. (2014). *Elliott Review into the Integrity and Assurance of Food Supply Networks – Final Report*. 146.

Everything you need to know about cocoa. (2019). *UTZ*. https://utz.org/better-business-hub/sourcing-sustainable-products/everything-need-know-cocoa/

Faberi, A., Marianella, R. M., Fuselli, F., Mantia, A. L., Ciardiello, F., Montesano, C., Mascini, M., Sergi, M., & Compagnone, D. (2014). Fatty acid composition and δ13C of bulk and individual fatty acids as marker for authenticating Italian PDO/PGI extra virgin olive oils by means of isotopic ratio mass spectrometry. *Journal of Mass Spectrometry*, *49*(9), 840–849. https://doi.org/10.1002/jms.3399

*Fairtrade and Child Labour*. (n.d.). Retrieved May 25, 2020, from https://fairtrade.org.nz/Fairtrade-in-Action/Fairtrade-and-Child-Labour

*FAQs*. (n.d.). OCHO. Retrieved May 20, 2020, from https://ocho.co.nz/pages/faqs

*Finishing with Perfection*. (n.d.). Lindt Chocolate World. Retrieved May 20, 2020, from https://www.chocolate.lindt.com/world-of-lindt/the-lindt-difference/the-lindt-differencefinishing-with-perfection/

*Food & Farming: Feeding a Growing Population*. (n.d.). Rainforest Alliance. Retrieved May 25, 2020, from https://www.rainforest-alliance.org/taxonomy/term/11

Francois, J. (2019). *UTZ How does the UTZ program address gender issues? - UTZ*. https://utz.org/better-business-hub/strengthening-your-reputation/international-womens-day-empowering-female-farmers/

Handley, L. L., & Raven, J. A. (1992). The use of natural abundance of nitrogen isotopes in plant physiology and ecology. *Plant, Cell and Environment*, *15*(9), 965–985. https://doi.org/10.1111/j.1365-3040.1992.tb01650.x

Ho, V. T. T., Zhao, J., & Fleet, G. (2014). Yeasts are essential for cocoa bean fermentation. *International Journal of Food Microbiology*, *174*, 72–87. https://doi.org/10.1016/j.ijfoodmicro.2013.12.014

Hoefs, J. (2015). *Stable isotope geochemistry* (Seventh edition.). Springer.

*How Chocolate Is Made*. (n.d.). Cadbury NZ. Retrieved December 14, 2018, from https://www.cadbury.co.nz/our-chocolate/how-chocolate-is-made/

*How to Make Chocolate From Scratch*. (n.d.). Instructables. Retrieved May 20, 2020, from https://www.instructables.com/id/How-to-Make-Chocolate-From-Scratch/

Hrastar, R., Petrišič, M. G., Ogrinc, N., & Košir, I. J. (2009). Fatty Acid and Stable Carbon Isotope Characterization of Camelina sativa Oil: Implications for Authentication. *Journal of Agricultural and Food Chemistry*, *57*(2), 579–585. https://doi.org/10.1021/jf8028144

Jahurul, M. H. A., Zaidul, I. S. M., Norulaini, N. A. N., Sahena, F., Jinap, S., Azmir, J., Sharif, K. M., & Omar, A. K. M. (2013). Cocoa butter fats and possibilities of substitution in food products concerning cocoa varieties, alternative sources, extraction methods, composition, and characteristics. *Journal of Food Engineering*, *117*(4), 467–476. https://doi.org/10.1016/j.jfoodeng.2012.09.024

Jinap, S., & Dimick, P. S. (1990). Acidic Characteristics of Fermented and Dried Cocoa Beans from Different Countries of Origin. *Journal of Food Science*, *55*(2), 547–550. https://doi.org/10.1111/j.1365-2621.1990.tb06806.x

Kandemir, T., Schuster, M. E., Senyshyn, A., Behrens, M., & Schlögl, R. (2013). The Haber–Bosch Process Revisited: On the Real Structure and Stability of “Ammonia Iron” under Working Conditions. *Angewandte Chemie International Edition*, *52*(48), 12723–12726. https://doi.org/10.1002/anie.201305812

Liu, Y., Zhang, X., Li, Y., & Wang, H. (2017). The application of compound-specific isotope analysis of fatty acids for traceability of sea cucumber (Apostichopus japonicus) in the coastal areas of China. *Journal of the Science of Food and Agriculture*, *97*(14), 4912–4921. https://doi.org/10.1002/jsfa.8367

Manly, B. F. J., & Navarro Alberto, J. A. (2016). *Multivariate Statistical Methods: A Primer, Fourth Edition*. CRC Press LLC. http://ebookcentral.proquest.com/lib/otago/detail.action?docID=4732224

Marshall, J. D., Brooks, J. R., & Lajtha, K. (2008). Sources of Variation in the Stable Isotopic Composition of Plants. In *Stable Isotopes in Ecology and Environmental Science* (pp. 22–60). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470691854.ch2

Meier-Augenstein, W. (2010). *Stable Isotope Forensics An Introduction to the Forensic Application of Stable Isotope Analysis* (1st ed.). Wiley.

Milton, R. D., Cai, R., Abdellaoui, S., Leech, D., De Lacey, A. L., Pita, M., & Minteer, S. D. (2017). Bioelectrochemical Haber–Bosch Process: An Ammonia-Producing H2/N2 Fuel Cell. *Angewandte Chemie International Edition*, *56*(10), 2680–2683. https://doi.org/10.1002/anie.201612500

Mubiru, E., Shrestha, K., Papastergiadis, A., & De Meulenaer, B. (2013). Improved gas chromatography-flame ionization detector analytical method for the analysis of epoxy fatty acids. *Journal of Chromatography A*, *1318*, 217–225. https://doi.org/10.1016/j.chroma.2013.10.025

Muccio, Z., & Jackson, G. P. (2009). Isotope ratio mass spectrometry. *Analyst*, *134*(2), 213–222. https://doi.org/10.1039/B808232D

O'Leary, M. H. (1981). Carbon isotope fractionation in plants. *Phytochemistry*, *20*(4), 553–567. https://doi.org/10.1016/0031-9422(81)85134-5

Perini, M., Bontempo, L., Ziller, L., Barbero, A., Caligiani, A., & Camin, F. (2016). Stable isotope composition of cocoa beans of different geographical origin. *Journal of Mass Spectrometry*, *51*(9), 684–689. https://doi.org/10.1002/jms.3833

*Rainforest Alliance Certified Cocoa*. (2016). Rainforest Alliance. https://www.rainforest-alliance.org/articles/rainforest-alliance-certified-cocoa

Rousseau, S. (2015). The role of organic and fair trade labels when choosing chocolate. *Food Quality and Preference*, *44*, 92–100. https://doi.org/10.1016/j.foodqual.2015.04.002

Sessions, A. L., Burgoyne, T. W., & Hayes, J. M. (2001). Determination of the H3 Factor in Hydrogen Isotope Ratio Monitoring Mass Spectrometry. *Analytical Chemistry*, *73*(2), 200–207. https://doi.org/10.1021/ac000488m

Sharp, Z. (2007). *Principles of stable isotope geochemistry* (1st ed.). Pearson Prentice Hall.

Silva, A. R. de A., Bioto, A. S., Efraim, P., & Queiroz, G. de C. (2017). Impact of sustainability labeling in the perception of sensory quality and purchase intention of chocolate consumers. *Journal of Cleaner Production*, *141*, 11–21. https://doi.org/10.1016/j.jclepro.2016.09.024

Spangenberg, J. E., & Dionisi, F. (2001). Characterization of Cocoa Butter and Cocoa Butter Equivalents by Bulk and Molecular Carbon Isotope Analyses: Implications for Vegetable Fat Quantification in Chocolate. *Journal of Agricultural and Food Chemistry*, *49*(9), 4271–4277. https://doi.org/10.1021/jf001509g

Torres-Moreno, M., Torrescasana, E., Salas-Salvadó, J., & Blanch, C. (2015). Nutritional composition and fatty acids profile in cocoa beans and chocolates with different geographical origin and processing conditions. *Food Chemistry*, *166*, 125–132. https://doi.org/10.1016/j.foodchem.2014.05.141

Traceability system. (n.d.). *UTZ*. Retrieved May 25, 2020, from https://utz.org/what-we-offer/traceability-system/traceability-system/

*Unique Roasting & Grinding*. (n.d.). Lindt Chocolate World. Retrieved May 20, 2020, from https://www.chocolate.lindt.com/world-of-lindt/the-lindt-difference/the-lindt-differenceunique-roasting-and-grinding/

Wakefield, J. (2013). *Application of stable light isotope ratios and trace element concentrations for the authentication of saffron: A thesis submitted for the degree of Master of Science, University of Otago, New Zealand*. Thesis MSc--University of Otago.

Walz, H. (2019). Why is certified cocoa better for the environment? *UTZ*. https://utz.org/better-business-hub/strengthening-your-reputation/sustainable-cocoa-better-environment/

*Why buy Fairtrade chocolate*. (n.d.). Retrieved May 25, 2020, from https://fairtrade.org.nz/Fairtrade-Products/Chocolate-cocoa/Why-buy-Fairtrade-chocolate

Young, M. E., & McCoy, A. W. (2016). Millennials and chocolate product ethics: Saying one thing and doing another. *Food Quality and Preference*, *49*, 42–53. https://doi.org/10.1016/j.foodqual.2015.11.014

Zhang, X., Han, D., Chen, X., Zhao, X., Cheng, J., & Liu, Y. (2019). Combined use of fatty acid profile and fatty acid δ13C fingerprinting for origin traceability of scallops (Patinopecten yessoensis, Chlamys farreri, and Argopecten irradians). *Food Chemistry*, *298*, 124966. https://doi.org/10.1016/j.foodchem.2019.124966

Ziegleder, G. (2009). Flavour Development in Cocoa and Chocolate. In *Industrial Chocolate Manufacture and Use* (pp. 169–191). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444301588.ch8

# Appendix 1

Bulk results of cocoa bean samples

| Sample I.D. | $\delta^2H$ | $\delta^{13}C$ | $\delta^{15}N$ | Origin |
|---|---|---|---|---|
| 81 | -117.352 | -28.2665 | 6.38 | Ghana |
| 81 | -116.884 | -28.3407 | 6.31 | Ghana |
| 81 | -114.466 | -28.5674 | 5.76 | Ghana |
| 86 | -106.193 | -27.8009 | 2.34 | Nicaragua |
| 86 | -105.012 | -27.1233 | 2.58 | Nicaragua |
| 86 | -102.367 | -27.7646 | 2.60 | Nicaragua |
| 87 | -113.803 | -26.5315 | 4.58 | Nicaragua |
| 87 | -114.172 | -26.6513 | 4.53 | Nicaragua |
| 87 | -113.435 | -26.6532 | 4.90 | Nicaragua |
| 89 | -113.239 | -25.9007 | 4.63 | Nicaragua |
| 89 | -116.339 | -26.7653 | 4.81 | Nicaragua |
| 89 | -113.731 | -26.277 | 4.42 | Nicaragua |
| 90 | | -26.5557 | 4.91 | Nicaragua |
| 90 | -99.898 | -26.8398 | 4.44 | Nicaragua |
| 90 | -100.255 | -26.4185 | 5.08 | Nicaragua |
| 88a | -108.291 | -27.9113 | 5.75 | Nicaragua |
| 88a | -111.162 | -27.8319 | 6.10 | Nicaragua |
| 88a | -105.354 | -27.6919 | 6.14 | Nicaragua |
| 88b | -105.495 | -27.8415 | 4.74 | Nicaragua |
| 88b | -99.2951 | -26.8531 | 5.07 | Nicaragua |
| 88b | -92.698 | -26.2318 | 4.95 | Nicaragua |
| 88c | | -26.2203 | 4.39 | Nicaragua |
| 88c | -116.092 | -27.3599 | 4.43 | Nicaragua |
| 88c | -114.699 | -27.1091 | 4.46 | Nicaragua |
| 2 | -150.606 | -29.4948 | 2.98 | PNG |
| 2 | -151.689 | -29.3816 | 3.09 | PNG |
| 2 | -152.433 | -29.9749 | 3.29 | PNG |
| 3 | -145.755 | -28.539 | 2.89 | PNG |
| 3 | -144.156 | -28.5619 | 2.87 | PNG |
| 3 | -143.614 | -28.4423 | 2.68 | PNG |
| 4 | -155.536 | -30.2234 | 0.80 | PNG |
| 4 | -154.991 | -29.3713 | 0.86 | PNG |
| 4 | -153.872 | -29.5499 | 0.92 | PNG |
| 5.1 | -170.196 | -29.3642 | 2.28 | PNG |
| 5.1 | -171.179 | -29.2601 | 2.11 | PNG |
| 5.1 | -171.945 | -29.5739 | 2.28 | PNG |
| 5.2 | -152.687 | -28.6065 | 3.70 | PNG |
| 5.2 | -156.75 | -29.8712 | 3.91 | PNG |
| 5.2 | -161.606 | -29.5299 | 3.85 | PNG |
| 5.3 | -168.372 | -28.3486 | 3.96 | PNG |
| 5.3 | -166.896 | -29.0983 | 4.12 | PNG |
| 5.3 | -168.111 | -28.3953 | 3.79 | PNG |

| | | | | |
|---|---|---|---|---|
| 6 | -146.84 | -31.2279 | 1.07 | PNG |
| 6 | -143.104 | -30.4605 | 1.12 | PNG |
| 6 | -144.287 | -30.635 | 1.34 | PNG |
| 7 | -156.343 | -29.7758 | 1.49 | PNG |
| 7 | -157.631 | -30.5589 | 1.97 | PNG |
| 7 | -155.285 | -30.3668 | 2.09 | PNG |
| 3a | -156.11 | -28.5009 | | PNG |
| 3a | -154.19 | -28.189 | | PNG |
| 3a | -158.105 | -28.6014 | | PNG |
| 3b | -148.839 | -28.8176 | | PNG |
| 3b | -151.041 | -28.5138 | | PNG |
| 3b | -152.252 | -28.0634 | | PNG |
| 3c | -152.605 | -29.5024 | | PNG |
| 3c | -152.74 | -29.1033 | | PNG |
| 3c | -150.588 | -28.7832 | | PNG |
| 4a | -146.556 | -30.39 | | PNG |
| 4a | -147.72 | -29.8474 | | PNG |
| 4a | -144.263 | -29.6341 | | PNG |
| 5a | -144.376 | -28.9719 | | PNG |
| 5a | -144.501 | -28.4351 | | PNG |
| 5a | -144.595 | -29.3168 | | PNG |
| 6a | -148.263 | -30.1204 | | PNG |
| 6a | -149.171 | -29.8337 | | PNG |
| 6a | -144.788 | -28.2459 | | PNG |
| 7a | | -29.0932 | | PNG |
| 7a | -137.424 | -28.7617 | | PNG |
| 7a | -138.997 | -28.4644 | | PNG |
| 82a | -136.89 | -29.7616 | 4.02 | Samoa |
| 82a | -140.266 | -30.344 | 2.98 | Samoa |
| 82a | -136.477 | -30.5647 | 2.86 | Samoa |
| 82b | -138.576 | -28.7717 | | Samoa |
| 82b | -138.299 | -28.8951 | | Samoa |
| 82b | -140.149 | -27.9943 | | Samoa |
| 82c | -137.209 | -29.7714 | | Samoa |
| 82c | | -28.8893 | | Samoa |
| 82c | -134.717 | -28.5457 | | Samoa |
| 8 | -131.978 | -31.7316 | 5.74 | Solomon Island |
| 8 | -128.538 | -31.4865 | 5.70 | Solomon Island |
| 8 | | -32.1262 | 5.55 | Solomon Island |
| 8b | -132.265 | -30.1846 | | Solomon Island |
| 8b | -134.71 | -29.1483 | | Solomon Island |
| 8b | | -30.1694 | | Solomon Island |
| 84 | -122.509 | -29.0692 | 7.15 | Vietnam |
| 84 | -120.674 | -28.0514 | 7.73 | Vietnam |
| 84 | -120.438 | -28.8195 | 7.42 | Vietnam |
| 85 | -134.581 | -29.4367 | 1.75 | Vietnam |

| | | | | |
|---|---|---|---|---|
| 85 | -132.397 | -29.4065 | 2.02 | Vietnam |
| 85 | -132.373 | -29.0217 | 1.73 | Vietnam |
| 91 | -129.289 | -28.3877 | 1.10 | Vietnam |
| 91 | -135.504 | -29.249 | 2.02 | Vietnam |
| 91 | -131.908 | -28.0066 | 1.27 | Vietnam |
| 83a | -122.505 | -26.0362 | 4.79 | Vietnam |
| 83a | -120.968 | -26.3587 | 5.33 | Vietnam |
| 83a | -121.631 | -25.3856 | 4.40 | Vietnam |
| 83b | -127.846 | -30.3935 | 3.73 | Vietnam |
| 83b | -127.577 | -30.5651 | 3.48 | Vietnam |
| 83b | -127.937 | -30.8671 | 3.72 | Vietnam |
| 83c | | -27.7734 | 3.46 | Vietnam |
| 83c | -110.056 | -27.6539 | 3.93 | Vietnam |
| 83c | -110.479 | -27.1727 | 4.05 | Vietnam |

# Appendix 2

Fatty acid results of cocoa bean and chocolate samples

| | Carbon | | | Hydrogen | | | |
|---|---|---|---|---|---|---|---|
| Sample | C16 | C18 | C18-1 | C16 | C18 | C18-1 | Origin |
| 81a 1 | -32.7809 | -32.5433 | -28.8774 | -135.833 | -136.558 | -160.992 | Ghana |
| 81a 2 | -33.5116 | -32.8873 | -29.5296 | -136.846 | -142.445 | -162.355 | Ghana |
| 81a 3 | -33.0635 | -33.0014 | -28.6198 | -135.015 | -138.086 | -147.175 | Ghana |
| 81b 1 | -33.4855 | -32.0261 | -29.6891 | -125.175 | -121.477 | -141.977 | Ghana |
| 81b 2 | -34.9086 | -33.0636 | -30.1534 | -129.092 | -129.839 | -150.941 | Ghana |
| 81b 3 | -33.9098 | -32.015 | -29.698 | -126.955 | -130.935 | -155.172 | Ghana |
| 81c 1 | -31.8901 | -32.0396 | -29.0703 | -133.737 | -137.361 | -156.798 | Ghana |
| 81c 2 | -33.5582 | -33.4338 | -28.4494 | -136.348 | -144.424 | -166.668 | Ghana |
| 81c 3 | -34.0739 | -33.4993 | -28.7187 | -132.577 | -136.56 | -145.965 | Ghana |
| 86 1 | -32.3724 | -33.3167 | -28.2405 | -127.073 | -131.852 | -137.285 | Nicaragua |
| 86 2 | -32.2788 | -33.0446 | -28.2334 | -124.463 | -130.312 | -155.445 | Nicaragua |
| 86 3 | -32.135 | -32.969 | -28.7826 | -129.31 | -130.655 | -154.112 | Nicaragua |
| 87 1 | -29.9882 | -29.9482 | -26.5937 | -124.928 | -130.449 | -151.58 | Nicaragua |
| 87 2 | -29.9048 | -29.6726 | -26.7867 | -124.271 | -128.741 | -147.355 | Nicaragua |
| 87 3 | -30.2547 | -30.4117 | -26.0776 | -119.771 | -132.113 | -153.465 | Nicaragua |
| 88a 1 | -32.7647 | -33.6215 | -28.8546 | -122.148 | -134.161 | -154.605 | Nicaragua |
| 88a 2 | -32.1101 | -32.9913 | -28.7807 | -126.15 | -130.919 | -150.516 | Nicaragua |
| 88a 3 | -32.7883 | -33.471 | -29.1137 | -125.201 | -131.231 | -149.941 | Nicaragua |
| 88b 1 | -31.5573 | -32.5733 | -28.786 | -117.015 | -128.292 | -146.105 | Nicaragua |
| 88b 2 | -31.2426 | -32.6215 | -28.7444 | -112.114 | -125.469 | -142.023 | Nicaragua |
| 88b 3 | -30.2519 | -31.6127 | -29.0023 | -108.573 | -125.112 | -141.052 | Nicaragua |
| 88c 1 | -30.5983 | -31.3633 | -27.3269 | -120.656 | -132.58 | -149.575 | Nicaragua |
| 88c 2 | -30.6493 | -31.3399 | -27.5473 | -119.759 | -126.69 | -139.853 | Nicaragua |
| 88c 3 | -29.5837 | -30.1628 | -27.8775 | -108.857 | -127.727 | -132.823 | Nicaragua |
| 89 1 | -30.8565 | -30.6969 | -27.3096 | -125.169 | -137.337 | -149.442 | Nicaragua |
| 89 2 | -29.8675 | -29.908 | -26.7023 | -114.584 | -132.692 | -139.947 | Nicaragua |
| 89 3 | -30.6981 | -30.7818 | -26.871 | -115.785 | -135.264 | -139.747 | Nicaragua |
| 90 1 | -80.169 | -114.773 | -118.755 | -80.169 | -114.773 | -118.755 | Nicaragua |
| 90 2 | -65.9494 | -109.966 | -94.0567 | -65.9494 | -109.966 | -94.0567 | Nicaragua |
| 90 3 | -75.6942 | -113.577 | -102.724 | -75.6942 | -113.577 | -102.724 | Nicaragua |
| 2a1 | -32.6823 | -32.5791 | -35.8809 | -175.576 | -164.279 | -183.442 | PNG |
| 2a2 | -34.2428 | -32.0154 | -34.8834 | -170.239 | -162.693 | -184.693 | PNG |
| 2a3 | -34.4861 | -31.9943 | -35.4576 | -167.563 | -161.868 | -174.776 | PNG |
| 3a1 | -33.2675 | -30.4078 | -34.3102 | -178.441 | -175.786 | -143.972 | PNG |
| 3a2 | | | | -176.217 | -173.943 | -184.789 | PNG |
| 4a1 | -36.0427 | -33.6019 | -36.1764 | -182.933 | -186.212 | -163.28 | PNG |
| 4a2 | -35.5869 | -33.5703 | -36.2619 | -186.141 | -183.664 | -157.485 | PNG |
| 4a3 | -35.7271 | -33.8806 | -37.328 | -189.772 | -187.759 | -190.687 | PNG |
| 5a1 | | | | -193.038 | -186.067 | -169.673 | PNG |
| 5a2 | | | | -192.055 | -185.872 | -167.278 | PNG |
| 5a3 | -34.177 | -30.7466 | -34.9393 | -194.698 | -188.839 | -174.023 | PNG |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 6a1 | -34.6168 | -31.8961 | -36.8182 | -181.478 | -175.958 | -172.573 | PNG |
| 6a2 | -34.6689 | -32.1263 | -34.6501 | -184.62 | -181.305 | -167.402 | PNG |
| 6a3 | | | | -180.188 | -177.624 | -157.019 | PNG |
| 7a1 | -33.4162 | -30.2379 | -34.2954 | -163.338 | -165.295 | -149.972 | PNG |
| 7a2 | -33.3248 | -30.9831 | -33.0055 | -160.88 | -163.058 | -144.948 | PNG |
| 7a3 | -33.395 | -30.545 | -33.4911 | -163.54 | -162.425 | -144.665 | PNG |
| 82 1 | -38.9074 | -38.3834 | -33.4519 | -164.856 | -170.251 | -178.616 | Samoa |
| 82 2 | -38.2718 | -37.2175 | -32.9435 | -168.272 | -174.413 | -198.224 | Samoa |
| 82 3 | -39.14 | -38.3013 | -33.4848 | -171.904 | -175.808 | -200.572 | Samoa |
| 8a1 | | | | -172.816 | -168.559 | -152.372 | Solomon Islands |
| 8a2 | | | | -164.472 | -165.121 | -153.585 | Solomon Islands |
| 8a3 | -35.3425 | -30.7593 | -35.4069 | -174.923 | -172.523 | -155.178 | Solomon Islands |
| 83a 1 | -33.2648 | -32.7581 | -28.0629 | -153.209 | -158.894 | -181.38 | Vietnam |
| 83a 2 | -34.2219 | -33.2508 | -28.4206 | -154.855 | -160.459 | -169.417 | Vietnam |
| 83b 1 | -32.8679 | -31.647 | -31.1825 | -139.551 | -142.23 | -169.855 | Vietnam |
| 83b 2 | -34.1246 | -33.9312 | -30.8216 | -140.11 | -144.437 | -170.513 | Vietnam |
| 83b 3 | -33.7416 | -32.9835 | -30.2211 | -141.015 | -145.738 | -171.284 | Vietnam |
| 83c 1 | -33.4927 | -33.6605 | -29.9966 | -159.005 | -169.963 | -194.718 | Vietnam |
| 83c 2 | -33.003 | -33.4566 | -29.9363 | -154.466 | -164.38 | -189.365 | Vietnam |
| 83c 3 | -29.2757 | -30.7586 | -31.8163 | -154.692 | -164.573 | -188.679 | Vietnam |
| 84 1 | -33.6857 | -33.4228 | -28.8269 | -135.687 | -141.222 | -159.471 | Vietnam |
| 84 2 | -33.6385 | -33.1697 | -29.1122 | -140.247 | -142.809 | -160.707 | Vietnam |
| 84 3 | -33.8046 | -33.3141 | -28.9167 | -139.063 | -141.5 | -160.38 | Vietnam |
| 85 1 | -33.9924 | -34.0136 | -30.2041 | -155.239 | -161.246 | -186.194 | Vietnam |
| 85 2 | -34.076 | -33.7732 | -30.3669 | -145.761 | -151.227 | -175.933 | Vietnam |
| 85 3 | -33.4454 | -33.5904 | -29.953 | -147.565 | -153.885 | -174.487 | Vietnam |
| 91 1 | -34.2906 | -34.2637 | -29.6298 | -118.046 | -147.028 | -136.053 | Vietnam |
| 91 2 | -34.5354 | -34.4409 | -29.4587 | -116.832 | -147.394 | -132.675 | Vietnam |
| 91 3 | -34.6672 | -34.5286 | -29.9106 | -110.486 | -142.853 | -125.339 | Vietnam |
| Choc 1 | -33.5818 | -33.7581 | -29.3195 | -161.863 | -146.347 | -186.253 | Ghana |
| Choc 1 | -34.2439 | -33.859 | -29.9165 | -162.206 | -154.416 | -191.278 | Ghana |
| Choc 1 | -33.5564 | -33.2381 | -29.9046 | -143.739 | -132.034 | -169.941 | Ghana |
| Choc 2 | -34.1281 | -33.8751 | -30.0614 | -143.602 | -136.89 | -138.726 | Ghana |
| Choc 2 | -33.4506 | -33.4236 | -30.3086 | -165.189 | -151.277 | -189.093 | Ghana |
| Choc 2 | -33.4184 | -33.5631 | -30.521 | -156.088 | -145.734 | -182.693 | Ghana |

# Appendix 3

Underivatised results of cocoa bean samples

| Sample | Peak 2 | Peak 1 | Origin |
|---|---|---|---|
| 81a | -45.926 | -35.866 | Ghana |
| 81b | -36.784 | -35.538 | Ghana |
| 86 | -35.508 | | Nicaragua |
| 87 | -32.237 | -34.412 | Nicaragua |
| 89 | -31.938 | -36.506 | Nicaragua |
| 90 | -31.264 | | Nicaragua |
| 88a | -32.842 | | Nicaragua |
| 88c | -32.443 | | Nicaragua |
| 2 | -39.605 | | PNG |
| 4 | -38.098 | | PNG |
| 5 | -37.786 | | PNG |
| 6 | -45.05 | -36.359 | PNG |
| 7 | -37.058 | | PNG |
| 3a | -40.945 | | PNG |
| 3c | -36.438 | -33.293 | PNG |
| 8 | -43.451 | | Solomon Islands |
| 84 | -34.261 | | Vietnam |
| 85 | -35.342 | -36.851 | Vietnam |
| 91 | -35.294 | | Vietnam |
| 83a | -35.321 | -35.029 | Vietnam |
| 83b | -38.064 | | Vietnam |
| 83c | -34.171 | -37.533 | Vietnam |